AD_______________

Award Number: W81XWH-04-1-0559

TITLE: Concurrent MR-NIR Imaging for Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Birsen Yazici

CONTRACTING ORGANIZATION: Rensselaer Polytechnic Institute
Troy, NY 12180

REPORT DATE: June 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20051018 062

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)* 01-06-2005	**2. REPORT TYPE** Annual	**3. DATES COVERED** *(From - To)* 1 Jun 2004 - 31 May 2005

4. TITLE AND SUBTITLE Concurrent MR-NIR Imaging for Breast Cancer Diagnosis	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER W81XWH-04-1-0559
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Birsen Yazici	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Rensselaer Polytechnic Institute Troy, NY 12180	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The primary objective of this research program is to investigate concurrent near infrared (NIR) optical and magnetic resonance (MR) imaging for breast cancer diagnosis. The NIR diffuse optical imaging offers novel criteria for cancer differentiation with the ability to measure (in vivo) oxygenation and vascularization state, the uptake and release of contrast agents and chromophore concentrations with high sensitivity. However, NIR diffuse optical tomography is inherently a low spatial resolution imaging modality due to diffuse nature of light photons. Alternatively, MRI provides high spatial resolution with excellent tissue discrimination, but has limited ability to monitor hemoglobin dynamics and other contrast mechanisms that optical imaging provides. Therefore, concurrent MRI-NIR optical imaging brings together the most advantageous aspects of the two imaging modalities for breast cancer diagnosis.

15. SUBJECT TERMS
Breast Cancer

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	130	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Cover..

SF 298...

Introduction.. 4

Body...6

Key Research Accomplishments... 33

Reportable Outcomes.. 34

Conclusions.. 35

References.. 36

Appendices.. 40

Concurrent MR-NIR Imaging for Breast Cancer Diagnosis

Birsen Yazici

I. INTRODUCTION

Near infrared (NIR) diffuse optical imaging provides quantitative functional information from breast tissue that can not be obtained by conventional radiological methods. NIR techniques can provide *in vivo* measurements of oxygenation and vascularization state, the uptake and release of molecular contrast agents and chromophore concentrations with high sensitivity. There is considerable evidence that tumor growth is dependent on angiogenesis [25]-[27], and that tumor aggressiveness can be assessed from its increased number of new vessels and reduced oxygenation state relative to normal breast tissue and benign breast lesions [28]-[30]. NIR diffuse optical tomographic (DOT) methods has the potential to characterize angiogenesis related vessel density as it measures the total hemoglobin concentration and provide the ability to differentiate between benign and malignant lesions based on oxygen saturation. Furthermore, NIR methods are non-ionizing, relatively inexpensive and can be made portable.

The diagnosis and management of cancer involves several stages where magnetic resonance (MR) plays a valuable and growing role. MRI of the breast is now a routine part of the clinical care in many centers [35]-[37]. Magnetic Resonance imaging (MRI) is indicated in patients with inconclusive clinical and/or mammographic examinations. Patients that may benefit include women with radiographically dense breasts, and high risk potential population [38]-[39]. MRI possesses less than 10% contrast for soft tissue pathology [40]. Gadolinium (Gd) enhanced MRI offers much better contrast and is specific for tumor vessel imaging. However, the signal in the Gd-MRI arises from the larger vessels as the contrast agent is flushed out of the vascular bed of the tumor [41]. In comparison, NIR measurements of absorption have extremely high contrast. It was reported that 5% change in vascular density as measured histologically in ductal carcinomas leads to approximately 300% contrast in NIR absorption coefficients [31]. Furthermore, there are studies suggesting that the kinetics of contrast enhanced optical spectroscopy provides information about the cellular spaces [34]. On the other hand, NIR DOT suffers from poor spatial resolution and as such, it is unlikely that NIR imaging will be a stand-alone screening method in the general population. Therefore, we believe that the concurrent MR and NIR imaging brings together the most advantageous aspects of the two imaging modalities (structural and functional). In the future, we envision that this multimodality imaging approach will lead to high resolution hemoglobin tomography and comprehensive quantitative functional tissue characterization to differentiate malignant and benign tumors.

In this project, the clinical studies are performed using the novel MR-NIR hybrid time-resolved spectroscopy (TRS) imager and fast Indocynine Green (ICG) enhanced spectroscopic imager developed by Dr. Chance, a Co-PI of this proposal, at the University of Pennsylvania (UPenn), Biophysics Department, Diffuse Optical Imaging and Spectroscopy Laboratory. The clinical study has been approved by the UPenn Institutional Review Board under. The project plans to leverage the expertise of Dr. Linda Nunes, M.D; Co-PI of this proposal, at Drexel School of Medicine, Radiologic Sciences Department, in systematic interpretation of MR breast architecture and kinetics for diagnosis. The PI and her collaborators at Upenn have been working on DOT image reconstruction problem and developed a number of approaches for DOT using a priori anatomical information in the past [42]. The PI also developed statistical tissue characterization methods for ultrasound breast cancer diagnosis [47]-[51]. Methods, tools and results developed in these studies are directly applicable to the proposed research.

The central hypothesis of this project is that the concurrent MR-NIR diffuse optical tomographic methods coupled with fast contrast enhanced NIR spectroscopic methods provide fundamentally new quantitative functional and structural information for breast cancer tumor characterization and detection. This new information can be obtained by novel modeling, analysis and data fusion methods from the tomographic, temporal and cellular-based contrast measurements, which exploit fast imaging techniques together with TRS tomographic methods. In this project, we investigate new methods for multi-modality high spatial resolution hemoglobin tomography, pharmacokinetic modeling of molecular contrast agents based on fast NIR spectroscopy and analysis of structural and functional information provided by MR and NIR imaging methods for breast cancer detection based on receiver operating characteristics methodology. Specific aims of the project are as follows:

- Aim 1: Utilize a priori anatomical information provided by MRI, to reconstruct 3D high resolution hemoglobin, water and lipid concentration, and oxygen saturation images directly from 6 wavelength time resolved optical measurements. Evaluate improvements in image reconstruction between that of stand-alone NIR and concurrent MR-NIR measurements using water and lipid images obtained from MRI.

- Aim 2: Develop a compartmentalized pharmacokinetic modeling of ICG, optical contrast agent, and extract quantitative parameters that can characterize tumor metabolism and angiogenesis. Compare ICG kinetics with the Gadolinium, MR contrast agent, kinetics and biopsy findings.
- Aim 3: Evaluate accuracy of breast cancer diagnosis based on the quantitative functional information extracted from stand-alone NIR system. This information includes hemoglobin, water and lipid concentration, optical scatter power and oxygen saturation images, and ICG pharmacokinetic parameters. Evaluate the added value of ICG kinetic parameters in breast cancer diagnosis.
- Aim 4: Combine NIR based breast cancer diagnosis features with the systematic MR breast architecture and kinetics interpretation model developed by Dr. Nunes, M.D, Co-PI of this proposal, to evaluate the sensitivity and specificity of concurrent MR-NIR imaging method. Compare results with that of stand-alone MR and NIR results.

In the following sections, we will provide detailed description of our current research in line with the statement of work (SOW) and the aims outlined above. For the period of June 1st, 2004 to May 31st 2005, SOW includes only the first two aims of the project.

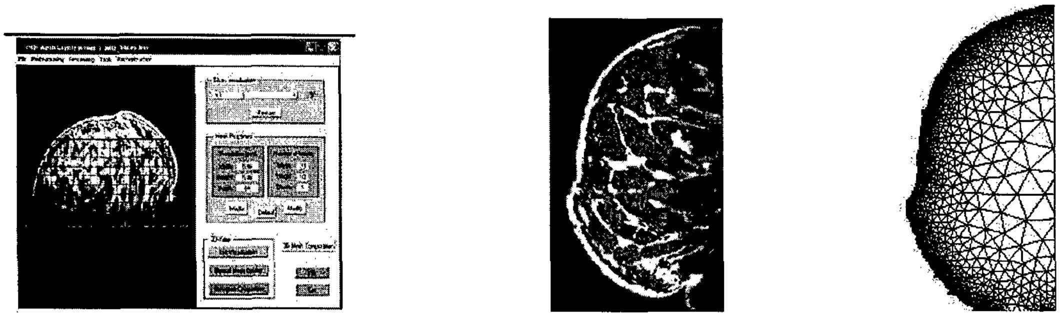
II. **BODY** : AIM 1- MR-GUIDED DIFFUSE OPTICAL TOMOGRAPHIC IMAGE RECONSTRUCTION

The SOW with regard to Aim 1 includes the following specific tasks:

- Task 1. Develop a portable toolbox to process MRI images to extract prior information that will be used to constrain the optical image reconstruction. This toolbox will perform 3D MR image segmentation using standard methods and extract the following information: Average optical properties of major breast tissue types, spatially varying optical grid structure. 0-2nd month. *Completed.*
- Task 2. Modify our existing optical image reconstruction algorithms that are based on perturbation model to incorporate diffusion equation in an iterative fashion. 2-8th month. *Completed.*
- Task 3. Perform direct reconstruction of hemoglobin concentration, oxygen saturation, water and lipid instead of optical absorption and scattering coefficients. 6-10th month. *Completed.*
- Task 4. Evaluate improvements in image reconstruction in terms of spatial resolution and quantitative accuracy between the stand-alone NIR and concurrent MR-NIR measurements using the water and lipid images obtained from MR images. 10-12th month. *Completed.*

A. *Aim1 - Task1:*

Figure 1 shows a snapshot of the MR a priori information processing toolbox that was developed by a team of two undergraduate students as a part of their senior design project under the PI's supervision. The code is programmed in C. The program can load a stack of clinical MR images and allows the user to interactively identify the coregistration fiducials mounted on the MR compression plates, perform 3D segmentation, interactively identify region of interest, define an adaptive mesh structure (2D), and perform standard 2D DOT image reconstruction (2D linear least-squares).



(a) A snapshot of the toolbox developed by the PI's undergraduate senior design team to perform preprocessing on MR images and basic DOT reconstruction.

(b) A slice of the segmented MR breast image.

(c) Adaptive mesh fit to the segmented data.

Fig. 1. The anatomical and optical images are shown on the left and right, respectively. Note the spatial mismatch between the two images.

B. *Aim1 - Task2:*

The outcome of our research regarding Aim 1, Task 2 are reported in the following publications: [52]-[55].

Two-level Domain Decomposition Methods (DDM) for Diffuse Optical Tomography (DOT) Two-level domain decomposition methods are introduced for the forward and the inverse problems of DOT. The proposed methods are TMODDM(Two-level Multiplicative Overlapping Domain Decomposition Method) for the forward solver and TMSDM(Two-level Multiplicative Space Decomposition Method). The convergence analysis and the estimation of computational costs are given. Numerical implementation shows that the proposed algorithm gives less mean square error and more clear image than the image produced by non-DDM. Furthermore DDM approach provides a framework to perform region of interest DOT image reconstruction based on the MR a priori anatomical information.

DOT is formulated by photon diffusion equation in the frequency domain. In the present approach, we have used the finite element method using piecewise bilinear element to compute photon density in the forward problem and the trust region method to compute optical coefficients which is a solution of nonlinear constraint minimization problem.

In DDM, the domain of interest is divided by overlapping subdomains as in Figure 2. Usually, the matrix solver for N number of nodes takes N^2 or N^3 order of computations. But at each subdomain the computation becomes $(N/d)^2$ or $(N/d)^3$ computations for each iterations. Thus if we use d number of parallel computers, the computation is the number of iteration times $1/(d^2)$ or $1/(d^3)$. Although, we use just 1 computer, the computation requires the number of iteration times $1/d$ or $1/(d^2)$.

Two-level method involves computing at the fine level with the aid of the coarse level reconstruction. Usually, the fine level computation decreases high frequency error, whereas the coarse level computation decreases low frequency components of the entire domain error. Using a two-level method for the forward problem, we can reduce the number of iterations. In 20×20 meshes, we get mean square error less than 10^{-5} for just 3 iterations. It is known that this number of iteration does not depend on the mesh size and subdomain size. Using many subdomains, we can get a more speed-up of the computation. In the inverse solver, the convergence is proved assuming the initial guess is sufficiently close to the target optical coefficients. Since the coarse level image approximates the low frequency of the target optical coefficients well, we have chosen coarse level correction before DDM step. Numerical experiments shows that these two-level DDM converges more fast than non-DDM in mean square error. The tumor image was more clear in two-level DDM than non-DDM.

The number of iteration obtained by using d subdomains is tabulated in Table 1. Here, N_n: The number of nodes, N_e: The number of elements d: The number of subdomains, M_F: The number of iteration for the forward solver, p: The order of computation with non-DDM. For nonsparse full matrix, $p = 3$, and for usual sparse matrix p is between 2 and 3.

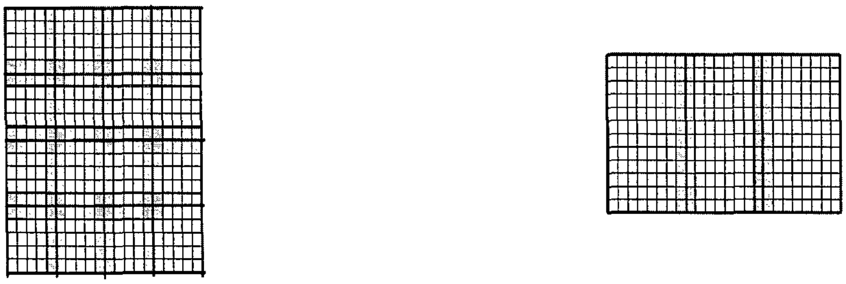
Thus if take more subdomains and use more parallel computers, we get much computation savings.

In the proposed method, if we know a priori information about the approximate location of tumor, using a prior information from secondary modality such as MRI, X-ray computed tomography (CT) or by a posteriori information from the coarse level image, we get better results by updating the coefficients

Method	Computational cost
non-DDM	$O(N_n^p) + O(N_e^p)$
TMSDM on 1 computer	$M_F d^{1-p} O(N_n^p) + d^{1-p} O(N_e^p)$
TMSDM on d computers	$M_F d^{1-p} O(N_n^p) + d^{-p} O(N_e^p)$
TMSDM on d^2 computers	$M_F d^{-p} O(N_n^p) + d^{-p} O(N_e^p)$

TABLE I

THE COMPARISON OF THE COMPUTATIONAL COSTS FOR THE NON-DDM METHOD, TMSDM FOR ONE, d, AND d^2 COMPUTERS.



(a) **(b)**

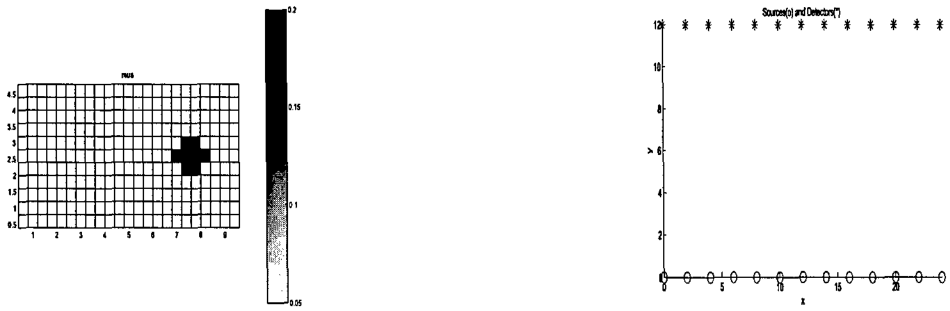
Fig. 2. (a) 4 × 4 domain decomposition with 20 × 20 mesh, (b) 3 × 1 domain decomposition with 24 × 12 mesh.

the special region of interest. Updating optical coefficients only in fourth subdomain in Figure 3(e), we have a better results than the image of TMSDM on all the subdomain in Figure 3 (d) or the image of non-DDM in Figure 3 (f).

1) Numerical results: The proposed method is implemented and compared with non-DDM in the following settings: $\mu_s' = 8cm^{-1}$, μ_a at background $= 0.05cm^{-1}$, μ_a at tumor $= 0.2cm^{-1}$, $\omega = 2\pi * 100MHz$, $c = 3 * 10^{10}$ cm/sec, $a = 1$.

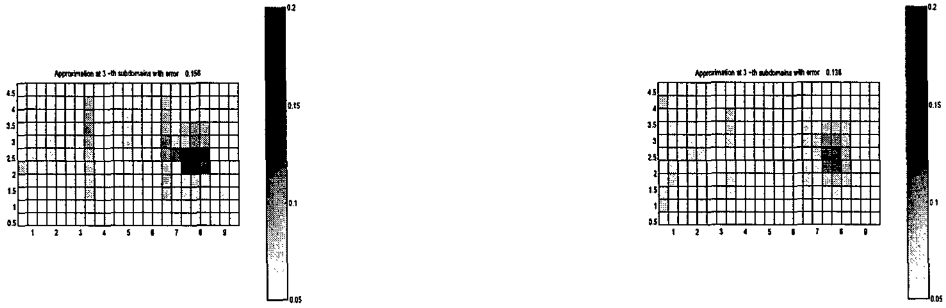
The construction of diffusion coefficient κ when μ_a is fixed is investigated in [53]. And the reconstruction of μ_a when μ_s' is fixed is implemented in Figures 3, and as in [54], [55].

The proposed method gives better result than non-DDM as in figure 3(d), as compared to figure 3(f). And we also get the results: Tikhonov parameters with $\alpha = 0.01$ gives better results than no regularization as in figure 3(d) as compared to figure 3(c). Detailed simulation studies can be found in [53], [54], [55].



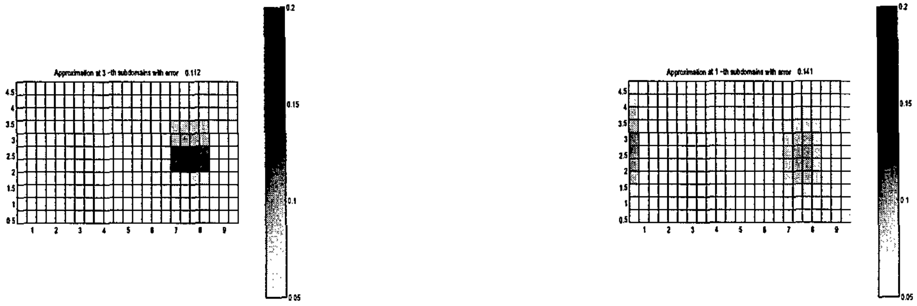
(a) The image of μ_a to be reconstructed. White region represent background tissue with $\mu_a = 0.05cm^{-1}$ and black region represent anomaly with $\mu_a = 0.2cm^{-1}$.

(b) The location of 20 sources and 20 detectors.



(c) The reconstructed image of the absorption coefficient using Algorithm TMSDM with $\alpha = 0$ resulting in L^2 norm error 0.0949.

(d) The reconstructed image of the absorption coefficient using Algorithm TMSDM with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.0824.



(e) The reconstructed image of the absorption coefficient using Algorithm TMSDM only on fourth subdomain with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.0755.

(f) The reconstructed image of the absorption coefficient not using Algorithm TMSDM with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.0926.

Fig. 3. The reconstruction of the absorption coefficient on $\Omega = [0,6] \times [0,6]$ with 2×2 subdomains, 20 detectors, 20 sources, and 400 pixels.

C. Aim1 - Tasks 3 and 4:

1) MR-Guided Hierarchical Bayesian Formulation of the Inverse DOT Problem: The outcome of our research regarding Aim 1, Task 3-4 are reported in the following publications: [2], [4], [42], [43], [44], [45], [46].

DOT poses a typical ill-posed inverse problem with limited number of measurements and inherently low spatial resolution. Both the low spatial resolution and ill-posedness of the inverse problem require the use of a priori information with high spatial resolution. In this context, we developed a hierarchical Bayesian approach to improve spatial resolution and quantitative accuracy by using *a priori* information provided by a secondary high resolution anatomical imaging modality, such as MR or X-ray. In such a dual imaging approach, while the correlation between optical and anatomical images may be high, it is not perfect. The proposed hierarchical Bayesian approach allows incorporation of partial *a priori* knowledge about the noise and unknown optical image models, thereby capturing the function-anatomy correlation effectively. In other words, the hierarchical Bayesian approach tackles with the fact that the correlation between the anatomical and optical images may be low. For example, there may be regions in the optical image that do not have any anatomical counterparts such that the tumor may be apparent in the optical image, but may not have a corresponding signature in the anatomical image. Furthermore, average optical coefficients extracted from anatomical images may be significantly different than the true optical coefficients of tissue. As a result, the assumption of strong optical-anatomy correlation may cause undesirable, erroneous bias in optical image reconstruction. Therefore, more flexible prior models are needed to properly represent optical-anatomy correlation. The hierarchical Bayesian framework affords such a flexibility in designing prior image and noise models. In the hierarchical Bayesian framework, one can formulate the inverse problem in multiple stages where each stage includes information about the unknown parameters of the preceding stage. The first stage of the hierarchy includes the data likelihood and the first stage of the image prior, which are comprised of statistical models for the noise and optical image, respectively. These models include parameters associated with noise and image statistics, which are not known precisely in practice. These unknown parameters are referred to as hyperparameters, which can be regarded as random variables. The succeeding stage of the hierarchical formulation incorporates *a priori* information about the hyperparameters in the form of prior distributions - so called hyperpriors - defined on the hyperparameters. The incorporation of the second stage concludes the design of the two-level hierarchical noise and image models.

In order to estimate the hyperparameters, we adapt the linear conjugate gradient (CG) algorithm to include a hyperparameter estimation step followed by an image update. In this context, we apply an iterative empirical Bayesian approach to estimate the hyperparameters, which in turn gives the maximum *a posteriori* (MAP) estimates of the hyperparameters at each CG iteration prior to the image update. Hence, the noise and image models are accommodated at each update of the hyperparameters along with the solution process. We refer to the following paper for further details [4] on hierarchical Bayesian formulation and estimation of hyperparameters.

In the following section, we present the optical image reconstruction results based on the developed hierarchical Bayesian approach in comparison to the maximum likelihood estimate of the same image. We consider two cases to show the effectiveness of the hierarchical Bayesian approach:

1) The tumor is present both anatomically and optically; thus perfect correlation exists between the anatomical and optical images.
2) The tumor is present optically, however it has no corresponding signature in the anatomical image. We show that the hierarchical Bayesian approach and the associated hyperparameter estimation effectively captures the inconsistency between the anatomical and optical images, thereby providing reliable optical image estimates.

2D Experiment with MR-simulated data:

We used the T1-weighted MR breast image from [1] to design a realistic optical breast model (figure 4). The MR breast image was segmented into parenchyma and adipose layers by applying a simple

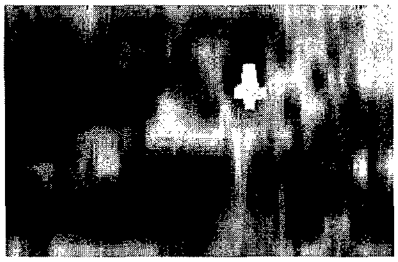
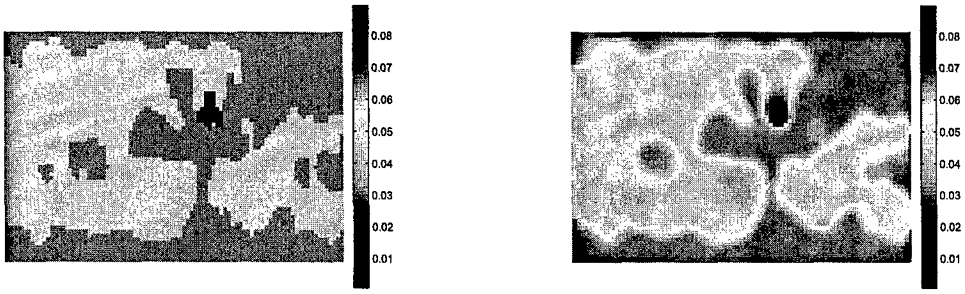


Fig. 4. The original MR breast image with an artificial tumor inserted.

thresholding algorithm with respect to the MR image intensity values. Next, a tumor corresponding to an infiltrating ductal carcinoma revealed by Gd-DTPA (Gadolinium-diethylenetriamine pentaacetic acid) enhancement was inserted (shown in figure 4 as well). Each sub-region was assigned an absorption value as indicated in [1] ($\mu_a^{adipose} = 0.03\text{cm}^{-1}, \mu_a^{parenchyma} = 0.06\ \text{cm}^{-1}$, $\mu_a^{tumor} = 0.09\ \text{cm}^{-1}$) to obtain an initial template (figure 5(b)). To simulate a corresponding optical image, zero mean Gaussian noise was added prior to filtering the image by a low-pass filter. The resulting optical image is shown in figure 5(b). Note the quantitative and spatial mismatch along the boundaries and especially within the tumor. The homogeneous diffusion coefficient of the medium was set to 0.042 cm. 9 frequencies ranging from 0 to 244 MHz were used to obtain 729 measurements with 9 sources and 9 detectors positioned along x-axis on opposite sides. The optical medium was uniformly discretized into 90 pixels along x-axis and 60 pixels along y-axis leading to a total of 5400 $1 \times 1\ \text{cm}^2$ pixels.



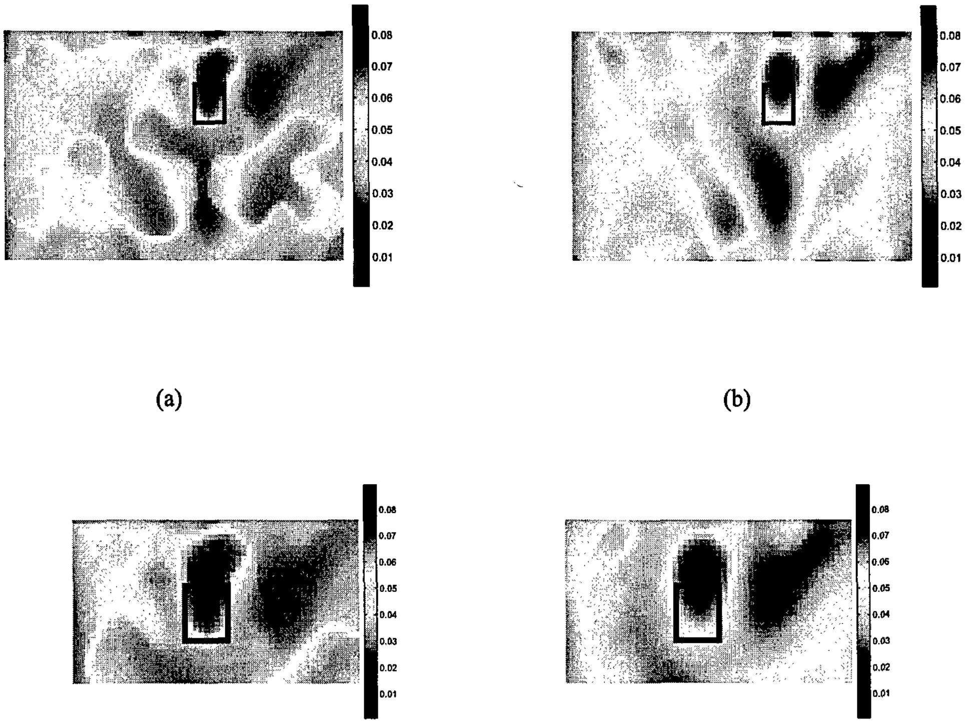
(a) (b)

Fig. 5. The anatomical and optical images are shown on the left and right, respectively. Note the spatial mismatch between the two images.

We performed two types of experiments to test the performance of the proposed hierarchical Bayesian approach for this problem:

(i) *Tumor present both anatomically and optically:* In this experiment, the template extracted from the anatomical image shown in figure 5(a) was used to design the hierarchical image prior. As a result, the optical image was segmented into three sub-images each of which corresponded to the labeled images in the anatomical image as shown in figure 5(a). In the design of the hyperprior defined on the mean (i.e. $p(\mu|\mathbf{C})$), values that significantly different than the actual mean of the sub-images were used. Thus, this experiment evaluates the robustness of the proposed method when the true statistics of the optical image are significantly different than the statistics extracted from the prior anatomical image.

The reconstructed image and the sub-image zoomed into the tumor region are shown in figures 6(a) and 6(c), respectively. For comparison, the maximum likelihood (ML) estimate of the image are shown in figures 6(b) and 6(d). The simulation results show that hierarchical Bayesian approach leads to qualitatively better results and resolves the tumor more accurately.



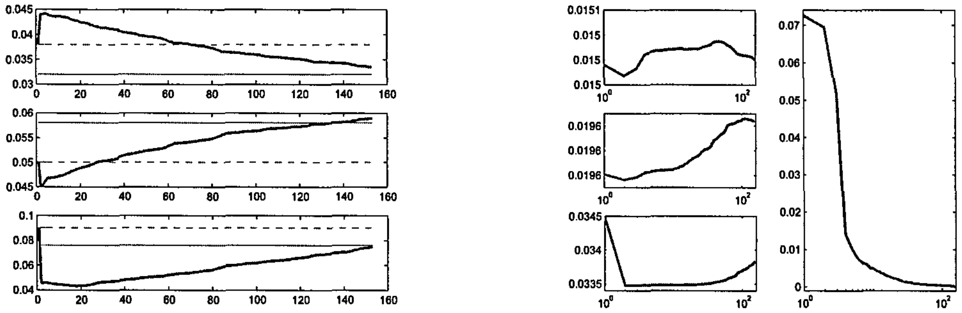
(a) (b)

(c) (d)

Fig. 6. The hierarchical Bayesian reconstruction of the optical image (a) using the anatomical template shown in figure 5(b) for the design of the hierarchical image model. Figure (c) shows the image that zooms into the tumor region in the optical image shown in (a). The maximum likelihood (ML) estimate of the entire image and the sub-image focusing the tumor region are shown in (b) and (d), respectively. The rectangular box in the figures shows the actual location of the tumor.

In figure 7, the estimates of the hyperparameters associated with the noise and image models are given as a function of the iteration number. Note that the mean value estimates for each sub-image converge to actual values, even though the corresponding assigned hyperparameters regarding the mean value deviate from the true average optical values by at least 15% (see table II). The experiment also demonstrates that the initialization of the hyperparameters does not have any effect on the performance of the estimation (figure 7).

(ii) *Tumor present optically but not anatomically:* In this experiment, we removed the tumor region from the template extracted from the prior anatomical image, but kept it in the optical image as shown in



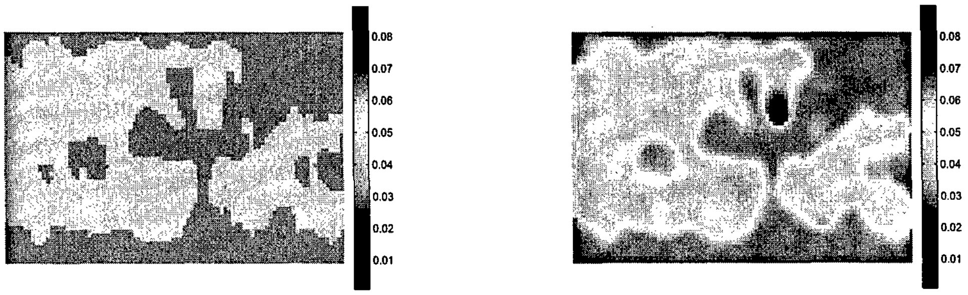
(a) The estimated mean values $\hat{\mu}_i$ for each sub-image vs iteration number are shown. The estimates for the parenchyma, adipose and tumor sub-images are given at the top, middle and bottom, respectively.

(b) The estimated standard deviation $\hat{\sigma}_i$ for each sub-image vs iteration number are shown on the left. The estimates for the parenchyma, adipose and tumor sub-images are given at the top, middle and bottom, respectively. The noise scale estimate $\hat{\lambda}$ vs iteration number is shown on the right.

Fig. 7. The mean value and the standard deviation estimates for each sub-image (sub-images are determined by the anatomical image) and noise scale estimate versus iteration number are shown. The thick solid line shows the estimated values. The constant solid line in (a) shows the actual mean value and the constant dashed line in (a) shows the assigned mean value $(\tilde{\mu}_i + \mu_{a0})$ used in the design of the hyperprior defined on the sub-image means. The actual mean values in these sub-regions are 0.032, 0.058 and 0.076, respectively.

figures 8(a) and 8(b). As a result, the optical image was segmented into two-sub images. The objective of this experiment is to evaluate how well the proposed method reconstructs optical tumors when they are not anatomically present.

The reconstructed images for this experiment are given in figures 9(a) and 9(c), respectively. The ML estimate of the image is given in figures 9(b) and 9(d). We observe that, even though there is a significant mismatch between the optical image and the anatomical counterpart in the tumor region, the hierarchical Bayesian formulation leads to a qualitatively better reconstruction than the ML approach, even around the tumor. Furthermore, the tumor is better localized as compared to the ML solution and is not biased towards the *a priori* anatomical image. The error in the localization of the tumor can be attributed to the source-detector geometry. The propagation of light along y-direction results in a smoothing effect on the optical image along y-direction. This effect is enhanced near source and detectors due to the behavior of the solution of the diffusion equation. The vertically smoothing effect can be observed in the ML



(a) (b)

Fig. 8. The anatomical template (a) and the original optical image (b). Note that the tumor is not anatomically present.

estimate of the image more apparently (figures 9(b) and 9(d)). The smoothing effect can be suppressed by incorporation of *a priori* information for the tumor as in case (i), where the tumor is better resolved (figure 6(c)). Further improvement can be achieved by employing sources and detectors positioned along y-axis as well as along x-axis.

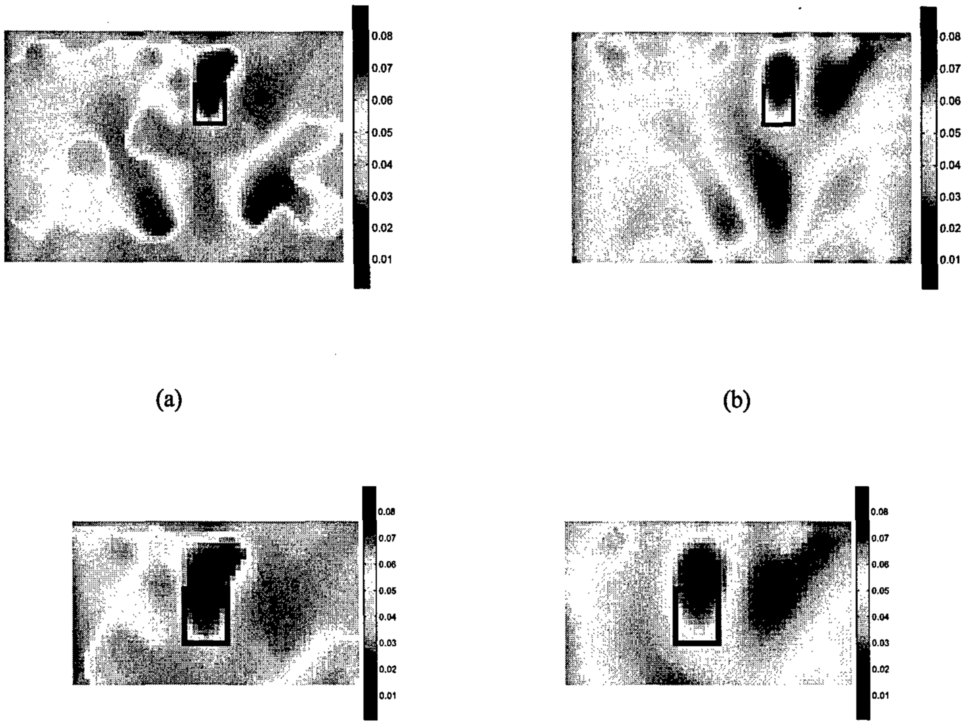
In figure 10, the average absorption value of each reconstructed sub-image vs iteration number is given. The sub-images correspond to those as indicated by the actual optical image shown in figure 5. Note that the mean value of the reconstructed image in the tumor region converges to the actual value even though the anatomical image asserts that no tumor exists.

The set of parameters used in the design of hyperpriors for these experiments and the actual mean of absorption values for each sub-image are shown in table II.

TABLE II

THE ACTUAL MEAN OF THE ABSORPTION VALUES IN EACH SUB-IMAGE AND THE PARAMETER SET USED IN THE INVERSE PROBLEM FORMULATIONS FOR THE MR-SIMULATED EXPERIMENTS I AND II. $\mu_{a0} = 0.0439$ FOR THIS EXPERIMENT. N/A STANDS FOR "NOT APPLICABLE".

			Sub-images	
		Parenchyma	**Adipose**	**Tumor**
The first experiment	$(\bar{\mu}_i + \mu_{a0}, \vartheta_i)$:	(0.038, 0.23)	(0.05, 0.3)	(0.09, 0.54)
	$(\mu_{\sigma_i}, \gamma_i)$:	(0.015, 0.228)	(0.02, 0.3)	(0.036, 0.54)
The second experiment	$(\bar{\mu}_i + \mu_{a0}, \vartheta_i)$:	(0.03, 0.18)	(0.06, 0.36)	N/A
	$(\mu_{\sigma_i}, \gamma_i)$:	(0.012, 0.18)	(0.024, 0.36)	N/A
	$(\bar{\mu}_a^{actual} + \mu_{a0})$:	0.032	0.058	0.076



(a) (b)

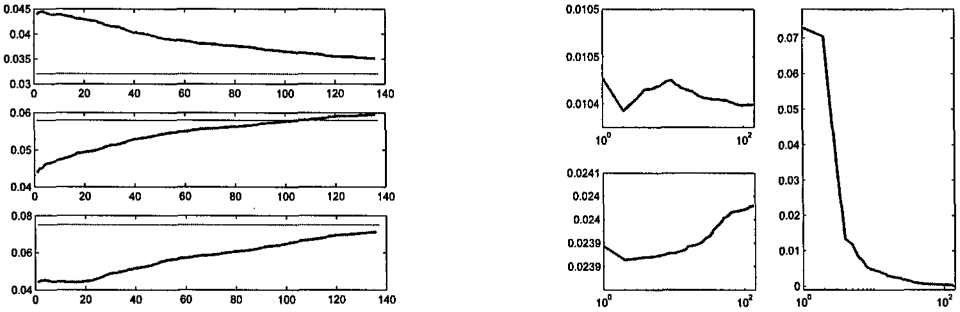
(c) (d)

Fig. 9. The hierarchical Bayesian reconstruction of the optical image (a) using the anatomical template shown in figure 8 for the design of the hierarchical image model. Figure (c) shows the image that zooms into the tumor region in the optical image shown in (a). The ML estimate of the entire image and the sub-image focusing the tumor region are shown in (b) and (d), respectively. The rectangular box in the figures shows the actual location of the tumor.

2) MR-Guided Direct Recovery of Blood Volume and Saturation: A detailed discussion of the research described in this section can be found in [2].

The estimation of the absorption coefficient at several wavelengths enables the provision of spatial maps of the targeted chromophores. In the case of the breast, the four chromophores of potential diagnostic interest are the oxy- and deoxy-haemoglobin, the water and the lipid. The concentrations of these breast constituents are linearly related to the absorption values.

Solving the linear system on a voxel basis for a spectral set of absorption distributions gives the functional maps required for diagnostic purposes, which can be termed as "indirect imaging". However, indirect imaging requires consecutive solutions of two linear systems, both of which are ill-posed. In order to overcome the burden and minimize the systematic errors due to the ill-conditioning of both inverse



(a) The average value ($\bar{\mu}_a$ in each reconstructed sub-image vs iteration number are shown. The estimates for the parenchyma, adipose and tumor sub-images are given at the top, middle and bottom, respectively. The domains of the sub-images correspond to the domains in the optical image shown in figure 8(a).

(b) The estimated standard deviation $\hat{\sigma}_i$ for each sub-image vs iteration number are shown on the left. The estimates for the parenchyma, adipose sub-images are given at the top and bottom, respectively. The sub-images correspond to the sub-images defined by the anatomical template shown in figure 8(b). The noise scale estimate $\hat{\lambda}$ vs iteration number is shown on the right.

Fig. 10. The mean values of the reconstructed sub-images vs iteration number (a). The estimated values of the standard deviation of the two sub-images and the noise scale λ are shown in (b).

problems, a method aiming at directly imaging the functional parameters has been proposed [32]. This method takes advantage of the linear relationship and formulates the inverse problem in terms unknown chromophore concentrations [2]. This new linear system is also poorly conditioned and great care should be taken during the pre-conditioning of this sensitivity matrix. In our case, we use an average column scheme [33]. This specific preconditioning scheme led to the most accurate reconstructions.

The functional parameters that we are interested in are the blood volume:

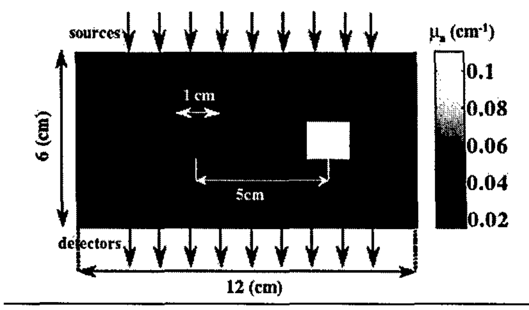
$$[BV] = [Hb] + [HbO_2] \tag{1}$$

and

$$[SaO_2] = \frac{[HbO_2]}{[Hb] + [HbO_2]} = \frac{[HbO_2]}{[BV]}. \tag{2}$$

Simulation Experiment Measurements were obtained by solving the frequency-domain diffusion equation with a finite difference approach (FDM). We restricted our simulations to a two-dimensional (2D) geometry for computational efficiency and to a continuous wave data set type. The slab thickness was

6 cm simulating a soft compressed breast. We placed nine sources on one side of the slab and nine detectors on the other side, both evenly stretched along 8 cm (see figure 11). Two square inclusions of 1 cm^2 were simulated in this model.



	Background	Left object	Right object
BV (mM)	0.02	0.06	0.12
SaO_2 (%)	70	65	50

	690	750	780	805	830	850
μ_s' (cm^{-1})	9.268	8.386	8.000	7.646	7.425	7.216
$\mu_a^{bkg.}$ (cm^{-1})	0.016	0.016	0.016	0.016	0.018	0.019
μ_a^{Left} (cm^{-1})	0.054	0.050	0.050	0.049	0.053	0.056
μ_a^{Right} (cm^{-1})	0.140	0.115	0.107	0.095	0.100	0.105

Fig. 11. Optical model used for the simulations (left).The optical properties displayed here correspond to $\lambda = 750nm$. Functional parameters used to define the optical properties to generate the synthetic measurements and optical parameters for the spectral set investigated. No $[H_2O]$ and [Li] constituents were used in these simulations for simplicity.

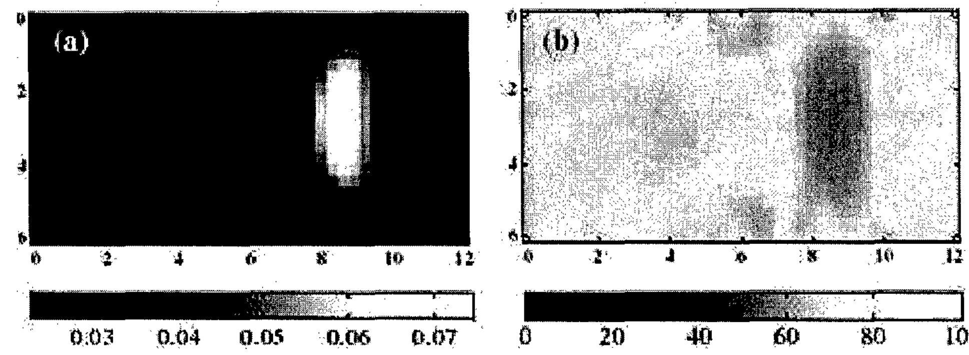


Fig. 12. (a) Blood volume (mM) reconstructions and (b) saturation estimates (%) in the case of classical indirect imaging.

First, the estimation of the functional parameters simulated is accomplished with a conjugate gradient descent (CGD) algorithm in the case of indirect imaging without using any kind of prior. The results are displayed in figure 12. Then reconstructions using the indirect imaging approach within the hierarchical Bayesian framework [2], [4] are provided in figure 13. First, an estimation of the absorption coefficients at the six wavelengths was performed using spatial *a priori* information and a conjugate gradient algorithm with the PolakRibiere method [4]. The assigned mean values for the absorption coefficients were the

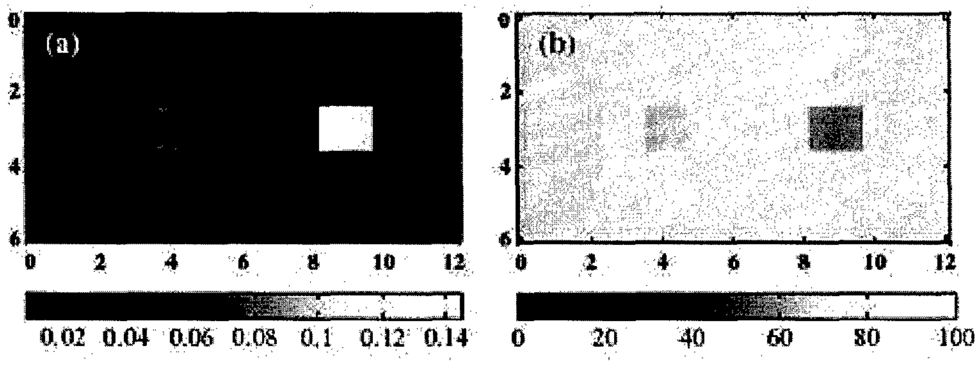


Fig. 13. O(a) Blood volume (mM) reconstructions and (b) saturation estimates (%) in the case of Bayesian indirect imaging.

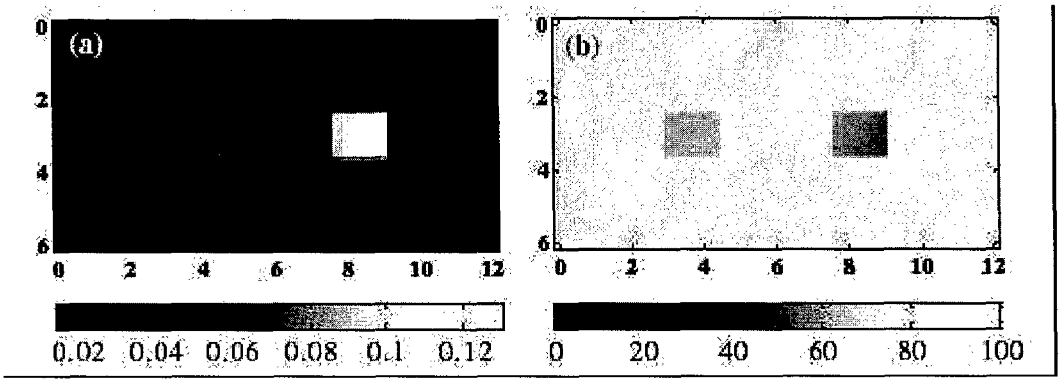


Fig. 14. (a) Blood volume (mM) reconstructions and (b) saturation estimates (Bayesian direct imaging. The assigned mean values correspond to the true concentration values with a 30% level of confidence.

simulated ones and 30% of the assigned mean values were used as the standard deviation. Then classical spectroscopy was performed on the resulting optical maps to solve for [HbO2] and [Hb] concentrations. In the second case, the results using physiological and spatial a priori information are depicted in figure 14. The assigned mean values for the chromophore concentrations were the exact ones with standard deviations 30% of the mean values. The functional quantitative values retrieved from each case and for the relevant region of interest are summarized in table 15.

In this work [2], we reported our first step towards incorporating physiological and spatial *a priori* information derived from MRI to assist DOT. Better estimates of the main functional parameters that are [BV] and [SaO2] were achieved.

	Background		Left object		Right object	
	[BV]	[SaO$_2$]	[BV]	[SaO$_2$]	[BV]	[SaO$_2$]
Indirect imaging	0.023	71	0.042	67	0.075	57
Bayesian indirect imaging	0.019	72.46	0.065	63.73	0.122	55.44
Bayesian direct imaging	0.020	69.95	0.061	65.27	0.127	50.02

Fig. 15. Functional parameters recovered with the three approaches and for the three different functional areas. The values proposed here correspond to the mean value of the entire ROI defined as the a priori spatial masks. [BV] in mM and [SaO2] in %.

III. **BODY** : AIM 2 - PHARMACOKINETIC MODELING OF ICG

Aim 2 of the SOW involves the following specific tasks:

- Task 1. Investigate compartmentalized pharmacokinetic models for Gd and ICG to model flow dynamics of ICG in extravascular extracellular space and plasma. 0-6th months. *Completed.*
- Task 2. Develop mathematical methods to extract parameters from ICG pharmacokinetic model to capture contrast agent flow kinetics. 6-9th months. *Completed.*
- Task 3. Perform simulation and in vivo clinical experiments to validate pharmacokinetic model. 9-15th months. *Ongoing.*
- Task 4. Compare ICG flow dynamics with the Gd flow dynamics of the breast MRI data. 15-18th months. *Ongoing.*

A. Aim 2 - Task 1 and 2:

The outcome of our research regarding Aim 2, Task 1 and 2 are reported in the following publications and presentations: [5], [6], [7], [8], [9].

In order to utilize a compartmentalized model for the ICG optical contrast agent dynamics, to extract parameters that can characterize tumor physiology, metabolism, and angiogenesis using fast temporal NIR spectroscopy we developed three different compartmental models namely, the four, three and two-compartment models, to model the pharmacokinetics of ICG in cancerous tumors.

The four-compartment model includes capillary region, interstitial fluid region (ISF), parenchymal cell region and intracellular binding site as compartments. Figure 16 illustrates the capillary and extracapillary space relevant to the four compartment model.The ICG, injected intravenously into the subject, can pass through from capillary into reversible binding site inside the cell through the interstitial fluid region and the parenchymal cell region [11], [12], [13]. Moreover, in advanced tumor stages, the leakiness around the tumor vessels is expected to increase, resulting in higher permeability rates during the transportation of ICG into the compartments.

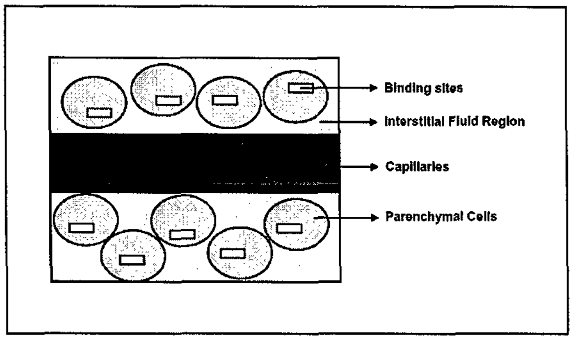


Fig. 16. A simple illustration of the capillary extracapillary structure.

In the three-compartment model, the parenchymal cell and intracellular binding site compartments are combined to form a single compartment called parenchymal cell. This amounts to the assumptions that the transition of ICG into the intracellular binding site is negligible as compared to the other compartments and therefore omitted from the model.

In the two-compartment model, the tumor region is assumed to be composed of two compartments; namely, the capillary region and ISF [10], [14], [15]. The transition of the ICG to the third and fourth compartments are assumed to be negligible. Therefore the last two compartments in the four compartment model is omitted. We consider transcapillary leakage to occur only at the tumor site. We also assume that a small perturbation of the global plasma concentration does not affect the bulk removal. The block diagrams of the compartmental transport and chemical model of ICG delivery for four, three, and two-compartment models are shown in figure 17 (a), (b), and (c), respectively.

The set of differential equations representing the ICG transition between the four compartments are given as follows:

The leakage into and the drainage out of plasma:

$$\frac{dC_p(t)}{dt} = k_b C_i(t) - k_a C_p(t) - k_{out} C_p(t). \tag{3}$$

The leakage into and the drainage out of ISF:

$$\frac{dC_i(t)}{dt} = k_a C_p(t) - k_b C_i(t) - k_c C_i(t) + k_d C_{pc}(t). \tag{4}$$

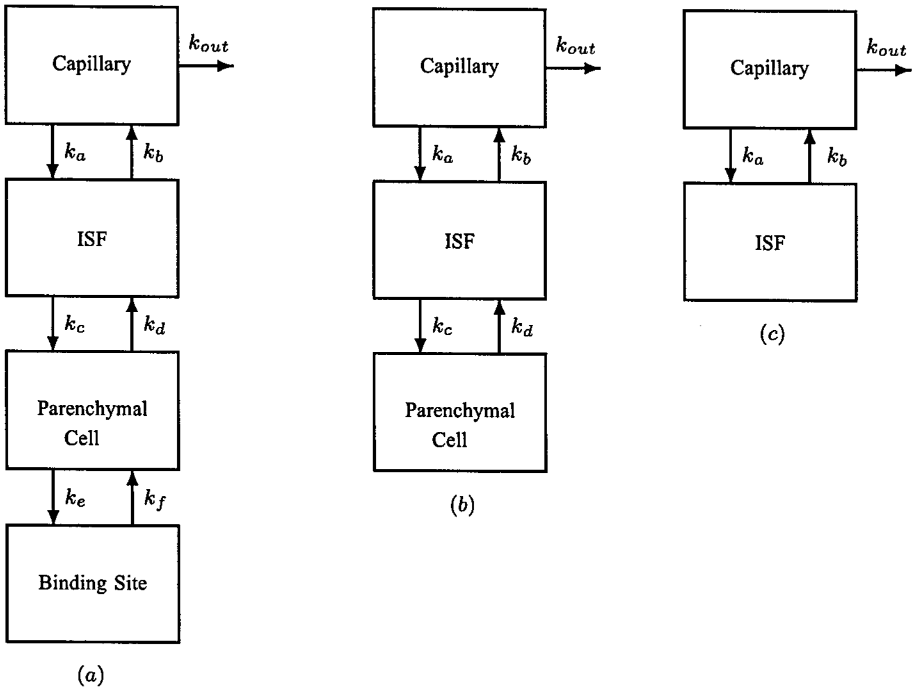


Fig. 17. Block diagrams of the (a) four-compartment, (b) three-compartment, (c) two-compartment models for the ICG pharmacokinetics.

The leakage into and the drainage out of parenchymal cell:

$$\frac{dC_{pc}(t)}{dt} = k_c C_i(t) - k_d C_{pc}(t) - k_e C_{pc}(t) + k_f C_b(t). \tag{5}$$

The leakage into and the drainage out of intracellular binding site:

$$\frac{dC_b(t)}{dt} = k_e C_{pc}(t) - k_f C_b(t). \tag{6}$$

Physiologically, the equilibrium constants are defined by the permeability surface area products given as $PS\rho$, where P is the capillary permeability constant, S is the capillary surface area, and ρ is the tissue density. k_{out} is proportional to the flow rate into and out of the capillary and k_a k_b, k_c , k_d, k_e, and k_f represent intra-tissue physiologic effects during ICG delivery from capillary to binding site.

The actual bulk ICG concentration in the tissue measured by NIR spectroscopy, $m(t)$, is a linear combination of the ICG concentrations in four different compartments.

$$m(t) = v_p C_p(t) + v_i C_i(t) + v_{pc} C_{pc}(t) + v_b C_b(t), \tag{7}$$

where v_p, v_i, v_{pc}, v_b, are volume fractions of plasma, ISF, parenchymal cell region and intracellular binding site, respectively.

The differential equations describing the transitions between compartments for three and two-compartments models are similar to the four-compartment model and described in detail in [5].

After compartmentalizing the region around the tumor we developed a mathematical estimation method to extract ICG kinetic rates. Using these rates one can differentiate between malignant and benign tumors and/or different tumor stages.

1) Kalman Filtring for the Estimation of Contrast Agent flow Kinetics: We introduce a systematic and robust approach to model and analyze ICG pharmacokinetics based on the extended Kalman filtering (EKF) framework. Kalman filter (KF) is an optimal recursive modeling and estimation method with numerous advantages in ICG pharmacokinetic modeling. These include:

- Effective modeling of multiple compartments, and multiple measurement systems governed by coupled ordinary differential equations, in the presence of measurement noise and uncertainties in the compartmental model dynamics.
- Simultaneous estimation of pharmacokinetic model parameters and ICG concentrations in each compartment, which is not accessible *in vivo* by means of NIR techniques.
- Recursive estimation of time-varying pharmacokinetic model parameters.
- Statistical validation of estimated concentrations and error bounds on the pharmacokinetic parameter estimates.
- Incorporation of available a priori information about the initial conditions of the permeability rates into the estimation procedure.
- Potential real-time monitoring of ICG pharmacokinetic parameters and ICG concentrations in different compartments due to the recursive nature of EKF estimation method.

The details of the extended Kalman filter for simultaneous estimation of pharmacokinetic model parameters and ICG concentrations in each compartment is explained in [5], [6].

For the optimal compartmental model order selection we adopted Bayesian information criteria (BIC). BIC is a well known information theoretic criteria, in which the optimal model order is selected by minimizing a cost function to avoid overfitting. The cost function depends on the number of observations, number of unknown parameters to be estimated and the maximum likelihood function. A detailed discussion on BIC can be found in [16], [17], [18]. In order to calculate the BIC for different compartmental models, we first derived a likelihood function for the extended Kalman filter. The derivation is based on the maximum likelihood estimation of the parameters in the Kalman filtering framework given as in [19], [20]. We then modified this likelihood function for the extended Kalman filter estimator for the joint estimation of compartmental model parameters and concentrations. The details of the derivation is provided in [5].

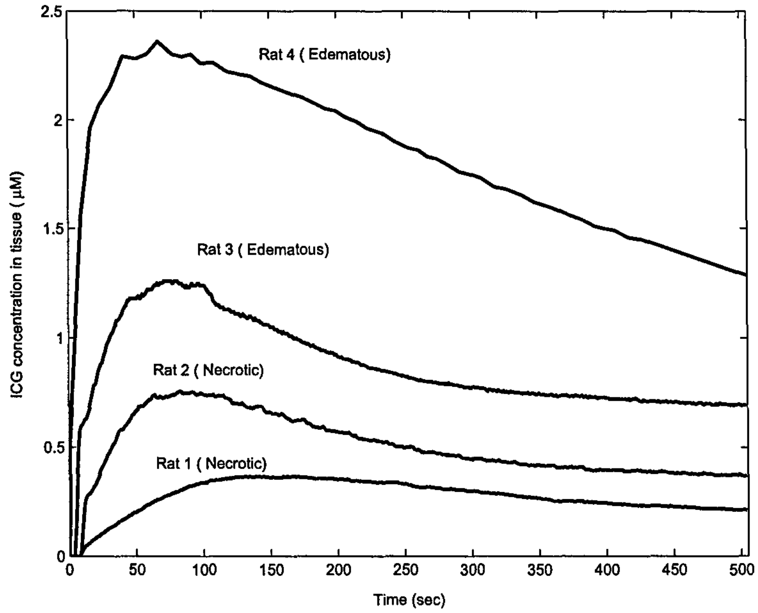


Fig. 18. ICG concentrations measured in tissue for four different rats.

B. Aim 2 - Task 3:

The outcome of our research regarding Aim 2, Task 3 are reported in the following publications and presentation: [5], [6], [7], [8], [9].

1) Fischer Rat Data Analysis: In order to validate the pharmacokinetic modeling of ICG kinetics we first used the ICG concentration data obtained from four Fischer rats with adenocarcinoma tumors. R3230ac adenocarcinoma cells were injected below the skin into four Fischer rats 3 weeks prior to measurements. The tumor size for the rats ranges in diameter from 5 to 30 mm. The ICG concentration data was collected with an MR-NIR imager. The configuration of the apparatus, the data collection procedure, and the details of the experimental approach have been reported in [10], [21], [22]. .

Figure 18 presents the ICG concentrations (μM) from four different rats. Tumors in Rat 1 and 2 are classified as necrotic because of their low tissue oxy-hemoglobin, low total hemoglobin, and low gadolinium-diethylene-triamine penta-acetic acid (Gd-DTPA) enhancement levels. Tumors in Rat 3 and 4 are classified as edematous due to their high water content. It can be observed from figure 18 that the necrotic cases display low peak ICG concentration values and slowly rising slopes unlike the edematous cases with high peak values and sharp rising slopes.

We estimated the pharmacokinetic rates for the four-, three- and two-compartment models. The results are given in tables III, IV, and V, respectively. The estimated pharmacokinetic rates for all compartmental

TABLE III

FOUR-COMPARTMENT MODEL:ESTIMATED PHARMACOKINETIC PARAMETERS USING EKF ALGORITHM

	$k_a^{(4)}$ $(sec^{-1}10^{-2})$	$k_b^{(4)}$ $(sec^{-1}10^{-2})$	$k_c^{(4)}$ $(sec^{-1}10^{-2})$	$k_d^{(4)}$ $(sec^{-1}10^{-2})$	$k_e^{(4)}$ $(sec^{-1}10^{-2})$	$k_f^{(4)}$ $(sec^{-1}10^{-2})$	$k_{out}^{(4)}$ $(sec^{-1}10^{-3})$
Rat 1 (Necrotic)	1.45±0.013	1.22±0.019	1.86±0.017	2.02±0.026	2.74±0.041	2.41±0.051	4.05±0.059
Rat 2 (Necrotic)	3.48±0.048	2.77±0.034	4.28±0.048	4.33±0.040	2.98±0.048	3.03±0.061	4.76±0.062
Rat 3 (Edematous)	4.94±0.052	5.16±0.067	4.22±0.052	4.13±0.067	4.14±0.070	4.27±0.078	5.39±0.085
Rat 4 (Edematous)	5.25±0.053	5.31±0.063	5.07±0.068	5.22±0.063	4.43±0.075	4.03±0.072	3.85±0.056

TABLE IV

THREE-COMPARTMENT MODEL: ESTIMATED PHARMACOKINETIC PARAMETERS USING EKF ALGORITHM

	$k_a^{(3)}$ $(sec^{-1}10^{-2})$	$k_b^{(3)}$ $(sec^{-1}10^{-2})$	$k_c^{(3)}$ $(sec^{-1}10^{-2})$	$k_d^{(3)}$ $(sec^{-1}10^{-2})$	$k_{out}^{(3)}$ $(sec^{-1}10^{-3})$
Rat 1 (Necrotic)	1.93±0.061	1.28±0.049	1.82±0.032	2.02±0.041	3.89±0.052
Rat 2 (Necrotic)	4.41±0.074	2.48±0.067	4.87±0.066	5.03±0.057	5.45±0.071
Rat 3 (Edematous)	4.71±0.085	3.88±0.077	4.95±0.059	4.68±0.050	4.42±0.040
Rat 4 (Edematous)	5.29±0.091	6.48±0.096	4.48±0.062	4.20±0.048	5.01±0.055

models indicate that the exchange rates between the capillary and the adjacent compartment, k_a^n, k_b^n, $n = 2, 3, 4$, are significantly different for the necrotic and edematous tissue. We observe that for the four- and three-compartment models, the estimated exchange rates between the ISF and parenchymal cell compartments, k_c^n, k_d^n, $n = 3, 4$, are comparable. Similarly, the estimated rate of drainage out of the plasma, k_{out}^n, $n = 2, 3, 4$, are consistent for all models.

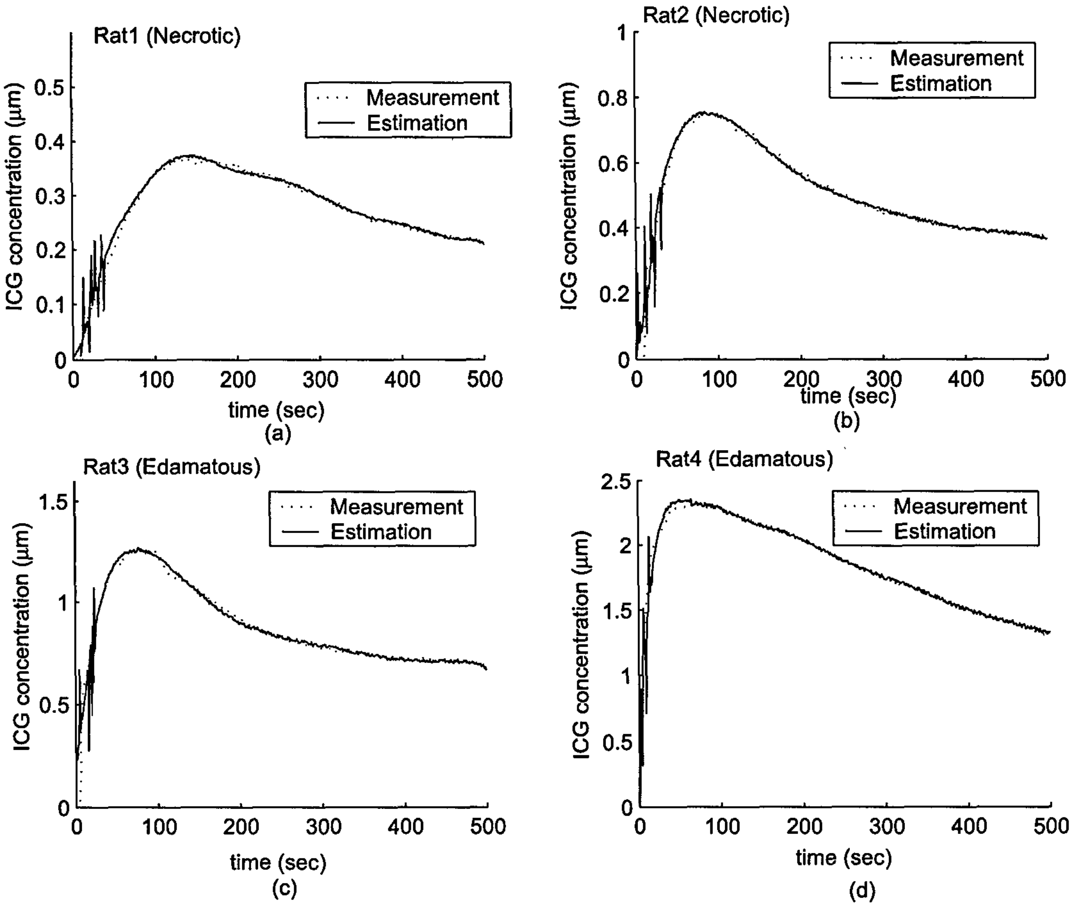
Based on the model parameter estimates, we computed the BIC values for each rat data to reveal overfitting. The BIC suggests that the two-compartment model is sufficient for all four measurement sets. We further analyze the goodness-of-fit of the compartmental models by means of the residual analysis. The details and results explaining why two-compartment model is a better statistical fit is given in [5]. Figure 19 shows the measured total concentration data and its 1-step ahead prediction based on the two-compartment model for each rat data. Clearly, there is a good agreement between the actual and the estimated measurements.

As we concluded that the two-compartment model provides the best statistical fit for the rat data, we investigated the estimated model parameters in more detail.

TABLE V

TWO-COMPARTMENT MODEL:ESTIMATED PHARMACOKINETIC PARAMETERS AND VOLUME FRACTIONS USING EKF ALGORITHM

	$k_a^{(2)}$ $(sec^{-1}10^{-2})$	$k_b^{(2)}$ $(sec^{-1}10^{-2})$	$k_{out}^{(2)}$ $(sec^{-1}10^{-3})$	$v_i^{(2)}$ (10^{-2})	$v_p^{(2)}$ (10^{-2})
Rat 1 (Necrotic)	2.47±0.043	1.06±0.052	4.61±0.073	21.8±1.92	1.41±0.053
Rat 2 (Necrotic)	3.54±0.082	2.98±0.086	4.83±0.092	25.4±3.49	2.42±0.088
Rat 3 (Edematous)	6.90±0.101	4.93±0.072	3.95±0.048	30.4±2.81	4.84±0.120
Rat 4 (Edematous)	8.40±0.114	7.77±0.091	4.02±0.068	53.0±4.73	7.03±0.321



(a) (b) (c) (d)

Fig. 19. ICG concentration measurement data and 1-step prediction of the measurements for four different rats.

In order to characterize the difference between these two tumors, we estimated the pharmacokinetic parameters k_a, k_b and k_{out}, and the volume fractions v_p and v_i for each rat in the two-compartment model. Table V tabulates the estimated parameters. The rate of leakage into the ISF from the capillary, k_a, range from 0.0247 to 0.0840 sec^{-1} and the rate of drainage out of the ISF and into the capillary, k_b, range from 0.0106 to 0.0777 sec^{-1}. Note that the permeability rates for the necrotic cases are lower than the ones observed for the edematous cases. Additionally, the estimated values for the pharmacokinetic rates are much higher than the normal tissue values due to the increased leakiness of the blood vessels around the tumor region. The estimated plasma volume fractions agrees with the values reported earlier [10]. These results confirm that v_p can be significantly large in tumors and that its magnitude varies with respect to the stage of the tumor. The estimated values of ISF volume fraction, v_i, range from 0.218 to 0.53, in agreement with 0.2 to 0.5 range reported earlier. Note that these results are valid only for the ICG pharmacokinetics in tumor cells R3230ac, adenocarcinoma and may not be generalized for other types of contrast agents or tumor types.

Figure 20 shows the estimated ICG concentrations in the plasma and the ISF compartments for the two-compartment model for Rats 1 to 4. The peak values of the plasma concentration, C_p, range from 2.72 μM to 4.28 μM. The absolute value of the concentrations may not be very useful. However, concentration of ICG in one compartment relative to the concentration in another compartment may provide useful information. We consider the ratio of the peak concentrations in the plasma and ISF as a potential parameter to discriminate different tumors. The peak C_p/C_i ratio for Rats 1 to 4 is 0.551, 0.593, 0.787, 1.151, respectively. This ratio is higher in edematous cases consistent with the fact that ICG-albumin leaks more into the ISF in advanced tumors. Additionally, the ICG concentration in plasma decays faster than the ICG concentration in ISF due to ICG elimination through the liver and kidneys.

2) Breast Data Analysis: In order to further investigate the pharmacokinetic modeling of ICG kinetics we are now working on the data collected from three different patients with different breast tumor types. First case, Case 1, is fibroadenoma, which corresponds to a mass estimated to be 1−2 cm in diameter and located around 6−7 o'clock by palpation within a breast of 9 cm diameter. Second case, Case 2, is adenocarcinoma corresponding to a tumor estimated to be 2−3 cm in diameter and located around 4−5 o'clock by palpation within a breast of 7.7 cm diameter. The third case, Case 3, is invasive ductal carcinoma, which corresponds to a mass estimated to be 3−4 cm in diameter and located around 6 o'clock. ICG was injected intravenously by bolus with a concentration of 0.25 mg per kg of body weight. Diagnostic information for tumors are obtained using biopsy results. Since biopsy modifies the

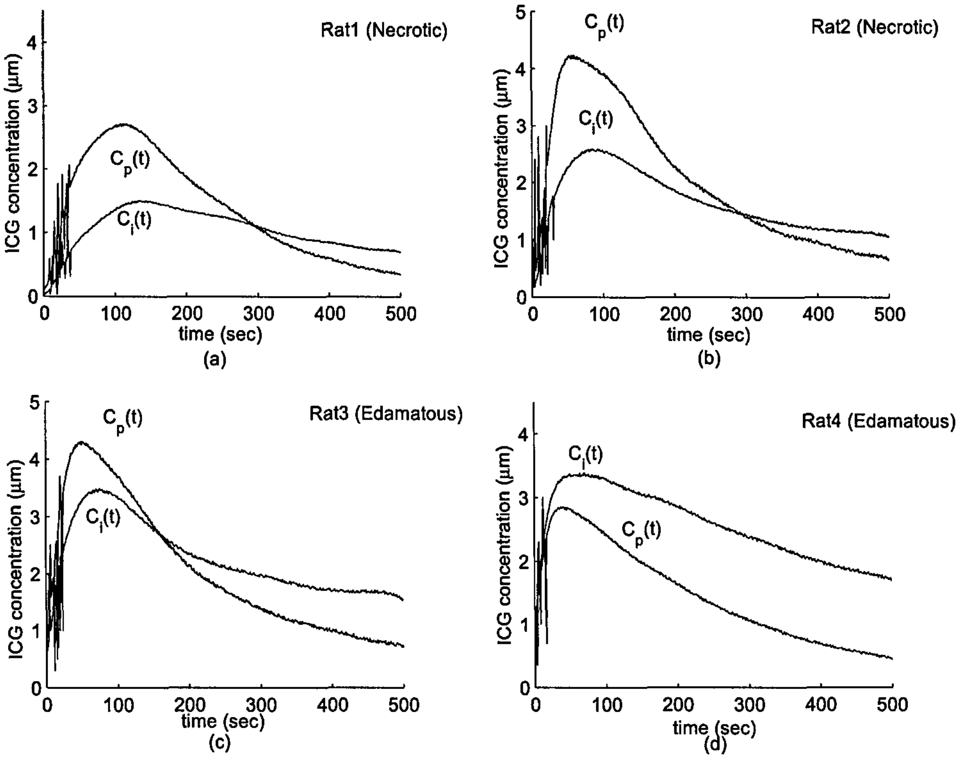


Fig. 20. ICG concentrations in plasma, $C_p(t)$ and ISF, $C_i(t)$, for four different rats. (a) Rat1, (b) Rat2, (c) Rat3, (d) Rat4.

blood volume and blood flow around the tumor region, measurements were made before the biopsy. Data acquisition started before injection of ICG and continued 10 minutes.

In this study, we make use of the data collected with a continuous wave (CW) NIR imaging apparatus. The apparatus has 16 light sources, namely, tungsten bulbs with less than 1 watt of output energy. They are located on a circular holder at an equal distance from each other with 22.5 degree apart. Sixteen detectors, namely, silicon photodiodes, are situated in the same plane. The breast is arranged in a pendular geometry with the source-detector probes gently touching its surface. Figure 21 illustrates the configuration of the apparatus and the configuration of the detectors and the sources in a circular plane. A band pass filter at 805nm, the absorption peak of ICG, is placed in front of the sources to select the desired wavelength. A set of data for one source is collected every 500 ms. The total time for a whole scan of the breast including 16 sources and 16 detectors is 8.8 seconds. The detectors use the same positions as the sources to collect the light originating from one source at a time. Only the signals from the farthest 11 detectors are used in the analysis. For example, when Source 1 is on, the data is collected using Detectors 4 to 14. The details of the data collection procedure and the apparatus is given in [23].

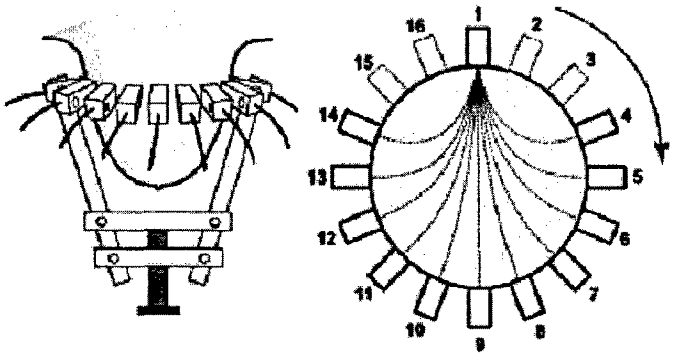


Fig. 21. The cut section of the 16 light source-detector device, holding a human breast inside. The diameter can be fitted easily. The 16 light source-detector combinations in each arm are located equal distance (11 degree apart), but when the device fits the breast, only the diameter chances.

Using the CW imager described above, for each patient, sufficient number of source detector readings were collected from different angles. Intes et al.[24] reconstructed the differential absorption images using these detector source readings for 46 different time instants. The details of the reconstruction algorithm is given in [24]. A sample set of reconstructed differential absorption images for Case 1, for 9 selected time instants are given in Figure 22.

In Figure 22 each one of 9 reconstructions consist of 649 voxels and each of these voxel values correspond to a differential absorption coefficient value. Using the the linear relationship between ICG concentration and absorption coefficient, we get ICG concentration maps from differential absorption mappings given in Figure 22. A sample set of ICG concentration maps for the selected time instants is illustrated in Figure 23. Here the concentration maps represents bulk ICG concentration in the tissue, not specifically in the plasma or the ISF.

ICG concentration curves for each voxel is traced using all reconstructions at different time intervals. For Case 1, we have 46 reconstructions. We have 649 ICG concentration curves representing 649 voxels and each of these curves has 46 data points representing 46 different time points. An example of the ICG concentration curve for the voxel where maximum uptake has occurred is given in Figure 24.

Using two compartmental model , we estimated pharmacokinetic parameters for each voxel using EKF algorithm. We then constructed 2-D permeability rate mappings using values of these parameters. 2-D maps for k_a, k_b for two-compartmental model are presented in Figures 25(a), 25(b).

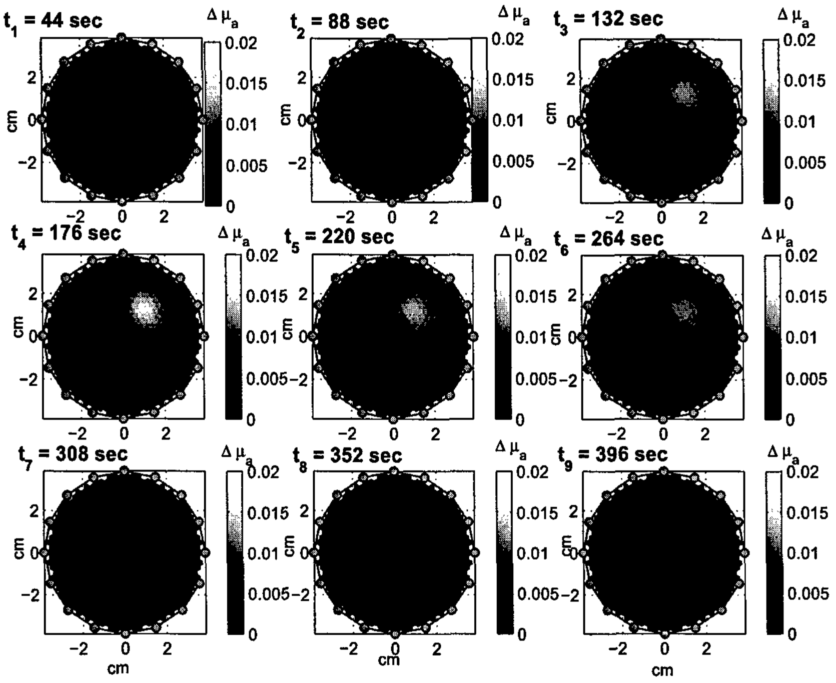


Fig. 22. Differential absorption reconstruction maps for the 9 different time instances.

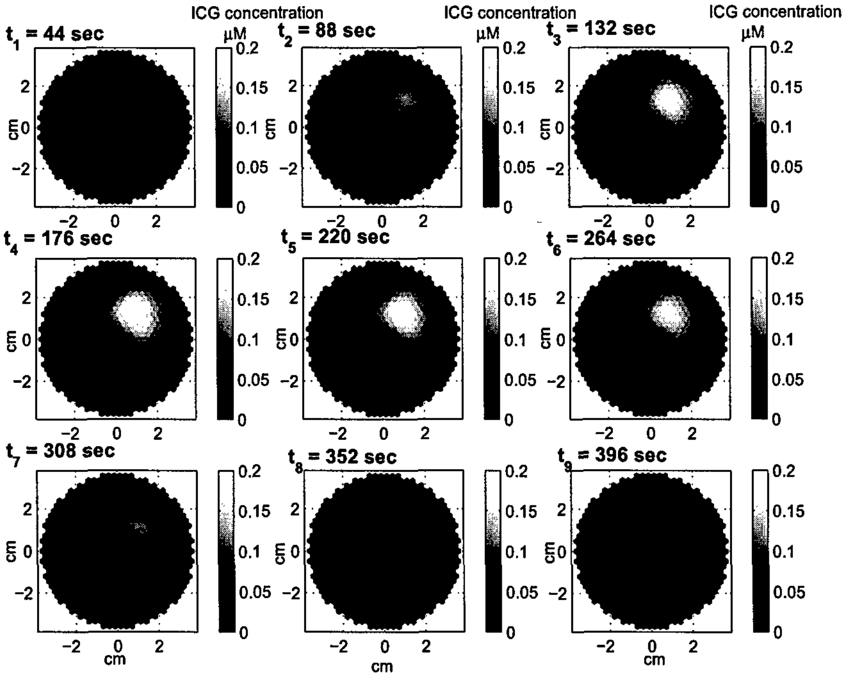


Fig. 23. 2-D ICG concentration mappings for selected time instances.

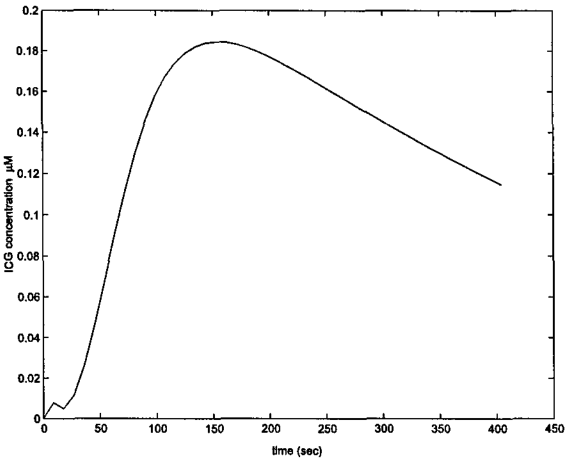
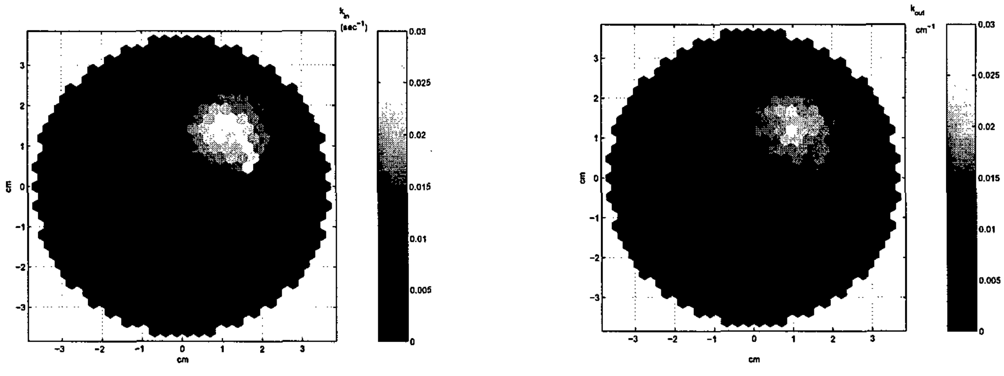


Fig. 24. ICG concentration curve for a specific voxel obtained using 46 different reconstructions at different time instants.


k_{in} (sec^{-1})
0.03
0.025
0.02
0.015
0.01
0.005
0
3
2
1
0
-1
-2
-3
cm
-3
-2
-1
0
1
2
3
cm
k_{out} cm^{-1}
0.03
0.025
0.02
0.015
0.01
0.005
0
3
2
1
0
-1
-2
-3
cm
-3
-2
-1
0
1
2
3
cm

(a) 2-D spatial map of the pharmacokinetic rates representing the drainage outside of the plasma

(b) 2-D spatial map of the pharmacokinetic rates representing the leakage into the plasma

Fig. 25. 2-D spatial mapping of pharmacokinetic rates representing the drainage outside of the plasma (a) and into the plasma (b)

We then analyzed if the kinetic rates are statistically different for inside and outside of the tumor region. The k_a and k_b values from inside and outside the tumor region are statistically different with a p-value of 0.0001. Using these results one may conclude that spatially resolved pharmacokinetic rate images can be used for breast cancer screening and diagnosis.

IV. KEY RESEARCH ACCOMPLISHMENTS

The SOW tasks defined for the first 12 months of the project have been completed. In particular,

1) An MR image analysis toolbox was developed. The toolbox performs image segmentation, extraction of anatomical prior information and co-registration of anatomical and optical images. (Aim 1, Task 1).
2) Two-level domain decomposition methods has been developed for diffuse optical tomography. The convergence of these methods are shown. The methods developed give smaller mean square error with less computational cost than the methods without domain decomposition techniques. The reconstruction algorithms can be made more efficient with parallel computing. If the tumor region of the optical coefficients is known by *a priori* anatomical information, the reconstruction algorithms become much more efficient and accurate by updating around the tumor region only. (Aim 1, Task 2)
3) The hierarchical Bayesian approach has been successfully applied for the recovery of absorption coefficients [4] and for the direct recovery of tissue chromophores [2] and has been proven to be effective in addressing the correlation between anatomical and functional images (Aim 1, Task 3 and 4)
4) Three different compartmental models have been developed for the flow kinetics of the optical contrast agent ICG. A robust and efficient method that can simultaneously estimate permeability rates and contrast agent kinetics in various compartments was developed. (Aim 2, Task 1 and 2)
5) The computational models and methods developed were applied to Fisher rat data carrying adenocarcinoma and shown that two compartment model is sufficient to characterize the kinetics of the ICG and to differentiate different tumor types. (Aim 2, Task 2 and 3)
6) A method to construct spatially varying pharmacokinetic rates is developed and is currently being applied to human breast data. (Aim 2, Task 3)

V. REPORTABLE OUTCOMES

The hierarchical Bayesian approach has been successfully applied for the recovery of absorption coefficients [4] and for the direct recovery of tissue chromophores [2].

Two-level multiplicative overlapping domain decomposition method (TMODDM) is applied to the linearized DOT [52]. With TMODDM as a forward solver, two-level multiplicative space decomposition method is adopted to the nonlinear DOT in [53], [54], [55]. The convergence and computational costs of the proposed algorithms are shown in [55]. The reconstruction of diffusion coefficient is shown in [53] and the reconstruction of absorption coefficients is shown in [54], [55].

Compartmental modeling of ICG pharmacokinetics using a two-compartment model for Fisher rat data has presented in [6]. Compartmental modeling of ICG pharmacokinetics using four, three, and two-compartment models and the best model order selection criteria has presented in [5] as a journal paper. A summary of this work has also presented in [6] and in [8]. Spatial mapping of pharmacokinetic rates using breast tumors have been presented in [9] Era of Hope meeting for the Department of Defense (DOD) Breast Cancer Research Program in June, 2005.

Complete list of reportable outcomes is given below:

1) Intes X, Maloux C, Guven M, Yazici B and Chance B 2004 Diffuse optical tomography with physiological and spatial a priori constraints *Phys. Med. Biol.* **49** N155–63
2) Guven M, Yazici B, Intes X and Chance B 2005 Diffuse optical tomography with a priori information *Phys. Med. Biol.* **50** 2837–58
3) M. Guven, B. Yazici, X. Intes, B. Chance, “3D Diffuse Optical Tomography with a priori Anatomical Information” SPIE-OSA Joint Proc. SPIE Vol. 5138, p. 268-280.
4) M. Guven, B. Yazici, X. Intes, B. Chance, 2003, “An Adaptive V-grid algorithm for Diffuse Optical Tomography,” International Conference on Image Processing, 2, 14–7.
5) M. Guven, B. Yazici, X. Intes, B. Chance, “Recursive least squares algorithm for optical diffusion tomography,” 2002, Bioengineering Conference Proceedings of the IEEE 28th Annual Northeast, 273–4.
6) B. Alacam, B. Yazici, X. Intes, B. Chance ” Extended Kalman Filtering for the Modeling and Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods”, *Transactions in IEEE Biomedical Engineering,* In Review.
7) B. Alacam, B. Yazici, X. Intes, B. Chance ”Extended Kalman Filtering Framework for the Modeling and Analysis of ICG Pharmacokinetics,” *Proc. Of 2005 SPIE Photonic West,* San Jose, California

USA, 22 - 27 January 2005, vol. 5693. pp. 17-27.

8) B. Alacam, B. Yazici, X. Intes, B. Chance "Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods," *Proc. of EMBS* – 27th *Anniversary Conference*, Shanghai, China, September 2005.
9) B. Yazici, K. Kwon, M. Guven, B. Alacam, " Concurrent MR-NIR Imaging for Breast Cancer Diagnosis", The Breast Cancer Imaging Network NTROI Annual Network Retreat and Workshop, Newport Beach, June 2-4, 2005.
10) B. Yazici, B. Alacam, X. Intes, N. Shoko, B. Chance, " Compartmental Modeling of ICG Pharmacokinetics for Breast Cancer Tumors" Era of Hope meeting for the Department of Defense (DOD) Breast Cancer Research Program, Philadelphia, June 8-11, 2005.
11) Il-Y. Son, M. Guven, B. Yazici, and X. Intes, 2004 A 2-level domain decomposition algorithm for inverse diffuse optical tomography, International Conference on Image Processing, *Proceedings of ICIP*, 5, 3315 - 3318.
12) K. Kwon, I.-Y. Son, and B. Yazici, 2005 Domain decomposition method for diffuse optical tomography, *Computational Imaging III*, edited by Charles A. Bouman and Eric L. Miller, *Proceedings of SPIE*, volume 5674, 64–75.
13) K. Kwon, I.-Y. Son, and B. Yazici, 2005 Two-level overlapping domain decomposition algorithm for a nonlinear inverse DOT problem, *Optical Tomography and Spectroscopy of Tissue VII*, edited by B. Chance, R. R. Alfano, B. J. Tromberg, M. Tamura, and E. M. Sevick-Muraca, *Proceedings of SPIE*, volume 5693, 459–468.
14) K. Kwon and B. Yazici, 2005 Two-level domain decomposition methods for diffuse optical tomography, in preparation.

VI. CONCLUSIONS

In the last 12 months, the research project has progressed as planned. Our research demonstrated that the computational requirements and accuracy of diffuse optical tomography can be significantly improved by the incorporation of *a priori* anatomical information provided by MRI. We developed new compartmental models and estimation methods to analyze the flow kinetics of the optical contrast agent ICG and applied our approach to Fisher rat data containing adenocarcinoma. Our results indicate that the flow kinetic parameters can be used to distinguish different tumor types. Currently, we are developing methods to build spatially resolved flow kinetic images using high temporal resolution NIR spectroscopy data. We are in the process of applying these methods to human breast data. These results were presented in 2005 Era of Hope Meeting, Philadelphia, PA. Presently, we expect the next 12 months of research to progress in line with the tasks outlined in the SOW.

REFERENCES

[1] Ntziachristos V, Chance B and Yodh A G 1999 Differential diffuse optical tomography *Optics Express* **5** 230–42

[2] Intes X, Maloux C, Guven M, Yazici B and Chance B 2004 Diffuse optical tomography with physiological and spatial a priori constraints *Phys. Med. Biol.* **49** N155–63

[3] Berger J O 1988 *Statistical Decision Theory and Bayesian Analysis* (Springer Verlag)

[4] Guven M, Yazici B, Intes X and Chance B 2005 Diffuse optical tomography with a priori information *Phys. Med. Biol.* **50** 2837–58

[5] B. Alacam, B. Yazici, X. Intes, B. Chance " Extended Kalman Filtering for the Modeling and Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods", *Transactions in IEEE Biomedical Engineering*, In Review.

[6] B. Alacam, B. Yazici, X. Intes, B. Chance "Extended Kalman Filtering Framework for the Modeling and Analysis of ICG Pharmacokinetics," *Proc. Of 2005 SPIE Photonic West,* San Jose, California USA, 22 - 27 January 2005, vol. 5693. pp. 17-27.

[7] B. Alacam, B. Yazici, X. Intes, B. Chance "Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods," *Proc. of EMBS* – 27th *Anniversary Conference*, Shanghai, China, September 2005.

[8] B. Yazici, K. Kwon, M. Guven, B. Alacam, " Concurrent MR-NIR Imaging for Breast Cancer Diagnosis", The Breast Cancer Imaging Network NTROI Annual Network Retreat and Workshop, Newport Beach, June 2-4, 2005.

[9] B. Yazici, B. Alacam, X. Intes, N. Shoko, B. Chance, " Compartmental Modeling of ICG Pharmacokinetics for Breast Cancer Tumors" Era of Hope meeting for the Department of Defense (DOD) Breast Cancer Research Program, Philadelphia, June 8-11, 2005.

[10] D.J. Cuccia, F. Bevilacqua, A. J. Durkin, S. Merritt, B. J. Tromberg, G. Gulsen, H. Yu, J. Wang, and O. Nalcioglu, "In vivo quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration," *Applied Optics*, vol. 42, no. 1, pp. 2940-2950, June 2003.

[11] P. Huang, X. Intes, B. Chance, S. Nioka, "Simulation of delivery of indocynanine green injected intravenously into the human subject for breast cancer detection," *Proceedings of SPIE*, vol. 4949, pp. 450-459, 2003.

[12] S. Mordon, J. M. Devoisselle, S. Soulie-Begu, T. Desmettre, "Indocyanine Green: Physicochemical Factors Effecting Its Flourescence in vivo," *Microvascular Research*, vol. 55, pp. 146-152, 1998.

[13] S. Fickweiler, R. M. Szeimies, W. Baumler, P. Steinbach, S. Karrer, A. E. Goetz, C. Abels, F. Hofstadler, M. Landthaler, "Indocyanine Green: Intracellular Uptake and Phototherapeutic effects in vitro," *Jour. Photochem. Photobio. B: Biology*, vol. 38, pp. 178-183, 1997.

[14] P. S. Tofts, DPhil, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp, H. B.W. Larsson, T. Lee, N. A. Mayr, G. J.M. Parker, R. E. Port, J. Taylor, and R. M. Weisskoff, "Estimating Kinetic Parameters From Dynamic Contrast-Enhanced T1-Weighted MRI of a Diffusable Tracer: Standardized Quantities and Symbols," *Jour. Mag. Res. Ima.*, vol. 10, pp. 223-232, 1999.

[15] P. S. Tofts, "Modeling tracer kinetics in dynamic Gd-DTPA MR imaging," *J. Magn. Reson. Imag.*, vol. 7, pp. 91-101, 1997.

[16] G. Schwarz, "Estimating the dimensions of a model," *Annals. of Statistics*, vol. 6, pp. 461-464,1978.

[17] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, Third edition, Prentice Hall, 1994.

[18] H. Akaike, "Likelihood and the Bayes Procedure," Bayesian Statistics, Univ. Press, Valencia, Spain, 1980.

[19] A. Harvey, *Time Series Models*, MIT Press, 1993.

[20] W.M. Sallas and D.A. Harville, "Noninformative Priors and Restricted Maximum Likelihood Estimation in the Kalman Filter", in J.C. Spall (Ed.), Bayesian Analysis of Time Series and Dynamic Models, New York: Marcel-Dekker, Inc., 1988.

[21] F. Bevilacqua, A. J. Berger, A. E. Cerussi, D. Jakubowski, and B. J. Tromberg, "Broadband absorption spectroscopy in turbid media by combined frequency-domain and steady-state methods," *Appl. Opt.*, vol. 39, pp. 6498-6507, 2000.

[22] D. J. Jakubowski, "Development of broadband quantitative tissue optical spectroscopy for the non-invasive characterization of breast disease," Beckman Laser Institute, University of California, Irvine, Irvine, Calif., 2002.

[23] S. Nioka, Y. Yung, M. Schnall, S. Zhao, S. Orel, C. Xie, B. Chance, and S. Solin, "Optical imaging of breast tumor by means of continuous waves," *Adv. Exp. Med. Biol.*, 411, 227-232, 1997.

[24] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, B. Chance, "*In vivo* continuous-wave optical breast imaging enhanced with Indocyanine Green," *Med. Phys.*, vol. 30-6, pp.1039-1047, June 2003.

[25] Folkman, J., Tumor angiogenesis: therapeutic implications. N. Engl. J. Med., 1971. 285(21): p. 1182-6.

[26] Folkman, J., What is the evidence that tumors are angiogenesis dependent? J. Natl. Cancer inst., 1990. 82(1): p. 4-6.

[27] Blood, C. H. and B. R. Zetter, Tumor interactions with the vasculature: angiogenesis and tumor metastasis. Biochem. Biophys. Acta, 1990. 1032(1): p. 89-118.

[28] Vaupel, P., et al., Current status of knowledge and critical issues in tumor oxygenation. Results from 25 years research in tumor pathophysiology. Adv. Exp. Med. Biol., 1998. 454: p. 591-602.

[29] Vaupel, P., et al., Oxygenation of human tumors: evaluation of tissue oxygen distribution in breast cancers by computerized O2 tension measurements. Cancer Res., 1991. 51(12): p. 3316-22.

[30] Runkel, S., et al., Oxygenations of mammary tumors as evaluated by ultrasound-guided computerized-pO2-histograpy. Adv. Exp. Med. Biol., 1994. 345: p. 451-8.

[31] Pogue, B. W., et al., Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: pilot results in the breast. Radiology, 2001. 218(1): 261–6.

[32] Durduran T, Giammarco J, Culver J, Choe R, Zubkov L, Holboke M, Intes X, Nioka S, Chance B and Yodh A 2001, Utilizing a priori spectral knowledge in diffuse optical tomography, *United Engineering Foundation Conf. Advances in Optics for Biotechnology, Medicine and Surgery VII.*

[33] Pei Y, Graber H and Barbour R 2001 Normalized constraint algorithm for minimizing inter-parameter crosstalk in DC optical tomography *Opt. Exp.* **9** 97-109

[34] Gurfinkel M et al 2000 Pharmokinetics of ICG and HPPH-car for the detection of normal and tumour tissue using fluorescence, near-infrared reflectance imaging: a case study Photochem. Photobiol. 72 94-102

[35] Harms S E. Breast magnetic resonance imaging. Seminars in Ultrasound, CT & MR 1998; 19:104-120.

[36] Harms SE, Flamig DP, Hesley KL, et al. MR imaging of the breast with rotating delivery of excitation off resonance: clinical experience with pathologic correlation. Radiology 1993; 187:493-501.

[37] Williams MB, Pisano ED, Schnall MD, Fajardo LL. Future directions in imaging of breast diseases. Radiology 1998; 206:297-300.

[38] Swift M, Morrell D, Massey RB, Chase CL. Incidence of cancer in 161 families affected by ataxia-telangiectasia. New England Journal of Medicine 1991; 325:1831-1836.

[39] Esserman L, Hylton N, Yassa L, et al. Utility of magnetic resonance imaging in the management of breast cancer: Evidence for improved preoperative staging. J Clin Oncol 1999; 17:110-119.

[40] Robinson, K. J., C. J. Kotre and K. Faulkner. "The use of contrast-detail test objects in the optimization of optical density in mammography." Br. J. Radiol. 68: 277-282, 1995.

[41] Moran, G. R. and F. S. Prato. "Modeling tissue contrast agent concentration: a solution to the tissue homogeneity model using a simulated arterial input function." Mag. Res. Med. 45:42-45, 2001.

[42] M. Guven, B. Yazici, X. Intes, B. Chance, "Diffuse optical tomography with a priori information derived from magnetic resonance images," Proceedings of Optical Tomography and Spectroscopy of Tissue V Conference at SPIE's Biomedical Optics 2003 meeting.

[43] M. Guven, B. Yazici, X. Intes, B. Chance, 2003, "An adaptive multigrid algorithm for region of interest diffuse optical tomography," Proceedings of International Conference on Image Processing, Volume 2, II-823–6 vol.3.

[44] M. Guven, B. Yazici, X. Intes, B. Chance, 2003, "An adaptive V-grid algorithm for diffuse optical tomography," Proceedings of IEEE 29th Annual Bioengineering Conference, 95–97

[45] M. Guven, B. Yazici, X. Intes, B. Chance, "Recursive least squares algorithm for optical diffusion tomography," 2002, Bioengineering Conference Proceedings of the IEEE 28th Annual Northeast, 273–4.

[46] B. Yazici, M. Guven, X. Intes, B. Chance, "MRI Assisted Diffuse Optical Tomography for Breast Imaging" accepted for publication in the 2003 International Microwave Symposium on June 13, 2003, Philadelphia, PA.

[47] B. Alacam, B. Yazici, N. Bilgutay, "Breast Tissue Characterization based on fractional differencing model of ultrasound RF echo," Proc. of SPIE Medical Imaging 2003 Proceedings , Vol. 5035, pp. 460–470, 2003.

[48] B. Alacam, B. Yazici, N. Bilgutay, "Breast cancer detection based on ultrasound B-scan texture analysis and patient age information," Proc. of 2003 IEEE 29th Annual Bioengineering Conference Proceedings, 98–99, March 2003.

[49] B. Alacam, B. Yazici, N. Bilgutay , "Breast Tissue Characterization based on Ultrasound RF echo modeling and tumor morphology," Proc. of EMBS-25th Anniversary Conference Proceedings, vol. 2, 1180–1183, September 2003.

[50] B. Alacam, B. Yazici, N. Bilgutay , "Breast cancer diagnosis based on ultrasound RF echo modeling and physician's level of confidence," IEEE International Ultrasonics Symposium, Hawaii, October 2003, vol. 2, 1203–1206

[51] B.Alacam, B.Yazici, N. Bilgutay, "Breast Tissue Characterization using FARMA Modeling of RF echo," Journal of Ultrasound in Medicine and Biology. Vol.30-10, October 2004, 1397–1407.

[52] Il-Y. Son, M. Guven, B. Yazici, and X. Intes, 2004 A 2-level domain decomposition algorithm for inverse diffuse optical tomography, International Conference on Image Processing, *Proceedings of ICIP*, 5, 3315 - 3318.

[53] K. Kwon, I.-Y. Son, and B. Yazici, 2005 Domain decomposition method for diffuse optical tomography, *Computational Imaging III* , edited by Charles A. Bouman and Eric L. Miller, *Proceedings of SPIE* , volume 5674, 64–75.

[54] K. Kwon, I.-Y. Son, and B. Yazici, 2005 Two-level overlapping domain decomposition algorithm for a nonlinear inverse

DOT problem, *Optical Tomography and Spectroscopy of Tissue VII*, edited by B. Chance, R. R. Alfano, B. J. Tromberg, M. Tamura, and E. M. Sevick-Muraca, *Proceedings of SPIE*, volume 5693, 459–468.

[55] K. Kwon and B. Yazici, 2005 Two-level domain decomposition methods for diffuse optical tomography, in preparation.

APPENDIX

(KEY PUBLICATIONS)

Two-level Domain Decomposition Methods for Diffuse Optical Tomography

Kiwoon Kwon† **and Birsen Yazici**† ‡

† Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute,
110 8th St., Troy, NY 12180, USA

Abstract. Diffuse optical tomography (DOT) in the near infrared involves the reconstruction of spatially varying optical properties of turbid medium from boundary measurements based on a forward model of photon propagation. Due to the highly non-linear nature of DOT, high quality image reconstruction is a computationally demanding problem that requires repeated solutions of both the forward and the inverse problems. Therefore, it is highly desirable to develop methods and algorithms that are computationally efficient. In this paper, domain decomposition methods are introduced to address the computational complexity of the DOT problem. A two-level multiplicative overlapping domain decomposition method for the forward problem and a two-level multiplicative space decomposition method for the inverse problem are developed. We show the local convergence of the inverse solver and derive the computational complexity of each method. The performance of the proposed methods in numerical simulations are demonstrated. The convergence analysis and computational cost analysis of both forward and inverse solvers are provided.

---

‡ corresponding author : yazici@ecse.rpi.edu

1. Introduction

Diffuse optical image reconstruction based on the diffusion equation is a highly non-linear ill-posed problem that calls for the use of non-linear minimization methods with regularization to stabilize the solution [1].

DOT reconstruction is posed as an optimization problem involving two coupled steps, namely forward and inverse problems. Each consists of an iterative solver whose solutions are used as an input to the other solver. More precisely, the forward solver computes the photon density and its Jacobian with respect to the optical coefficients, and the inverse solver updates the optical coefficients based on the output of the forward step. The updated coefficients are then used in the forward solver to recompute the photon density and its Jacobian. As a result, the computational complexity of the DOT reconstruction using these approaches, quickly grows with the number of pixels and dimensions. Thus, real time computation of DOT requires numerical techniques to reduce the complexity of the problem.

In this paper, we investigate domain decomposition methods(DDM) to address the computational requirements of the DOT image reconstruction. Domain decomposition methods have been developed in the last two decades in the area of numerical solutions of partial differential equations motivated by the need for fast and efficient algorithms for solving large-scale, three-dimensional problems. A major advantage of the DDM is that they allow for distributed parallel numerical solvers on smaller subdomains, making the computation extremely efficient. We applied DDM to the forward problem formulated by the solution of an elliptic partial equation given optical coefficients and space decomposition method to the inverse problem formulated as a minimization problem given boundary measurements. Badea et al. introduced the term space decomposition to describe the application of DDM approach to the constrained convex minimization problems arising from variational inequalities and provided conditions under which DDM based approaches lead to the convergence [4]-[6],[23]-[25]. DDM begins by partitioning the domain into two or more subdomains as in Figure 1. The inverse problem is then divided into subproblems on each of the subdomains.

Among many kinds of DDM techniques we have considered the multiplicative overlapping domain decomposition method (MODDM) for the forward problem and multiplicative space decomposition (MSDM) for the inverse problem. These methods are called one-level methods. The proposed methods in this paper are two-level MODDM (TMODDM) and two-level MSDM (TMSDM) for forward and inverse problems, respectively.

Since the photon density function is rapidly varying around the point source, the contribution from the sources may be neglected in the subdomains that are far apart from the sources. Thus the one-level MODDM is not sufficient to approximate the photon density function for the point source, since the subdomain correction step is confined to each subdomain. Therefore we propose an algorithm based on TMODDM for the forward solver[21]. The coarse level correction step applied to the entire domain decreases low frequency errors that are not handled well at the subdomain correction step. Therefore TMODDM decreases the overall error more efficiently than MODDM.

The use of two-level method is extended to the inverse problem. The convergence of MSDM is shown in [5] under the some conditions for the objective function F. Among the conditions, the strong convexity condition for F, that is the positive-definiteness of the Hessian of F, can be shown only locally. Thus to prove the convergence of MSDM, we require that the optical coefficients for each iteration step and at each subdomain must be sufficiently close to the target optical coefficient. In other words, a good initial guess for the target optical coefficient is required. Therefore, we have chosen upsampled coarse level image as an initial guess for Algorithm MSDM assuming it is sufficiently close at least at low frequency.

The paper is organized as follows: In the next section, the forward solver, based on the photon diffusion equation model and minimization formulation of the inverse DOT, is described. Algorithms applying domain decomposition techniques to the forward and inverse steps of the minimization formulation are presented in Section 3. Section 4 presents numerical simulations and Section 5 summarizes our results and conclusions. The paper includes an appendix for the proof of the local convergence of TMSDM algorithm.

2. Diffuse Optical Tomography

2.1. Photon Diffusion Equation in Frequency Domain

Propagation of light in biological tissues is well-modelled by the photon diffusion equation with the Robin boundary condition. In frequency domain photon diffusion equation is given as follows:

$$\begin{aligned} &-\nabla \cdot (\kappa \nabla \Phi) + (\mu_a + \frac{i\omega}{c})\Phi = q && \text{in } \Omega \\ &\Phi + 2a\kappa \frac{\partial \Phi}{\partial \nu} = 0 && \text{on } \partial\Omega \end{aligned} \tag{2.1}$$

where Ω is a Lipschitz domain in $\mathbb{R}^n, n = 2, 3$, $\partial\Omega$ is its boundary, ν is the unit outward normal vector on the boundary, Φ is the photon density, q is a source term, and μ_a, μ_s', and $\kappa = \frac{1}{3(\mu_a+\mu_s')}$ are the absorption, reduced scattering, and diffusion coefficients, respectively. The constant R, is determined by the refraction on the boundary.

The unique identification of the optical coefficients μ_a and κ in (2.1) when Dirichlet-to-Neumann map is given (or infinite sources and infinite detectors are given), can be easily shown by using the uniqueness results for isotropic case [22]. For the uniqueness of the optical coefficients when κ has anisotropic anomalous region contained in a known background, see [10, 11, 12].

On the boundary, we measure the following Rytov data [18]:

$$\Gamma = \log\left(-\kappa\frac{\partial\Phi}{\partial\nu}\right) \tag{2.2a}$$

$$= \log\left(\frac{1}{2a}\Phi\right). \tag{2.2b}$$

[2, 3] show numerical examples where the Rytov approximation provides much better images than the Born approximation.

In this paper, we assume that μ_s' is known and concentrate only on the reconstruction of the absorption coefficient μ_a. Thus, the Jacobian of Γ with respect to μ_a for given μ_s' as follows:

$$\frac{\partial\Gamma}{\partial\mu_a}(r) = -\frac{1}{2a\Phi}\int_\Omega[-3\kappa^2\nabla G(r,r')\nabla\overline{\Phi(r')} + G(r,r')\overline{\Phi(r')}]dr', \tag{2.3}$$

where $\overline{\Phi}$ is the complex conjugate of Φ.

2.2. Discretization and Finite Element Method

Suppose N_s sources $r_j, j = 1,\cdots,N_s$ and N_d detectors $r_{i+N_s}, i = 1,\cdots,N_d$ on boundary of Ω is given. Let Φ_j be the solution of (2.1) for point source $q^j(r) = \delta(r - r_j), j = 1,\cdots,N_s$. Define

$$\Gamma_{i,j} = \log(\Phi_j(r_{i+N_s})), \tag{2.4}$$

for i-th detector and j-th source.

Consider the finite element space with N_n $u_k, k = 1,\cdots,N_n$ bases. For piecewise bilinear element, N_n is the same as the number of nodes. Let the number of elements be N_e and denote the elements by $T_m, m = 1,\cdots,N_e$.

The finite element formulation for (2.1) in this finite element space for the j-th source q_j $(j = 1,\cdots,N_s)$ is as follows:

$$\left[K + C + \frac{1}{2a}A\right]\Phi_j = b^j, \tag{2.5}$$

where K, C, A are $N_n \times N_n$ matrices and b^j is a $N_n \times 1$ column vector given by

$$\begin{aligned} K_{kl} &= \int_\Omega \kappa\nabla u_k\nabla\overline{u_l} \\ C_{kl} &= \int_\Omega\left(\mu_a + i\frac{\omega}{c}\right)u_k\overline{u_l} \\ A_{kl} &= \int_{\partial\Omega} u_k\overline{u_l} \end{aligned}$$

$$b_l^j = \begin{cases} 1 & \text{if } l = j \\ 0 & \text{Otherwise,} \end{cases}$$

for $k, l = 1, \cdots, N_n$.

Assume that L and U are constants such that $0 < L < U$. Define a function space for μ_a as

$$\begin{aligned} V &= \{\mu_a \in L^2(\Omega) | L \leq |\mu_a(r)|_{r\in\Omega} \leq U\}, \\ V_{N_e} &= \{\mu_a \in V | \mu_a \text{ is constant at each } T_m, m = 1, \cdots, N_e\}. \end{aligned}$$

Define norms in V and V_{N_e} as follows:

$$\|x\|_V = \|x\|_{L^2(\Omega)} \text{ for } x \in V,$$

$$\|x\|_{V_{N_e}} = \sqrt{\sum_{m=1}^{N_e} \frac{x(m)^2}{|T_m|}} \text{ for } x \in V_{N_e},$$

where $x(m)$ is the constant at the element T_m and $|T_m|$ is the area of $T_m, m = 1, \cdots, N_e$.

Assume that $\kappa, \mu_a \in V_{Ne}$. Let $\Phi_j(k)$ be the value of Φ_j at k-th node point and $\kappa(m)$ be the value of κ at m-th element T_m. From the solution of (2.5), we obtain the boundary measurement data $\Gamma_{i,j}$. By discretizing (2.3), we get the value of the Jacobian of $\Gamma_{i,j}$ at the m-th element:

$$\begin{aligned} \frac{\partial \Gamma_{i,j}}{\partial \mu_a}(m) &= -\frac{1}{2a\Phi_j(i)} \sum_{k,l=1}^{N_n} \Phi_i(k)[-3\kappa(m)^2 V_m(k,l) + W_m(k,l)]\overline{\Phi_j(l)} \\ V_m(k,l) &= \int_{T_m} \nabla u_k \nabla \overline{u_l}, \\ W_m(k,l) &= \int_{T_m} u_k \overline{u_l}. \end{aligned} \tag{2.6}$$

2.3. *DOT as a nonlinear ill-posed optimization problem and the trust region method*

Given (2.1) and (2.6), DOT is formulated as the following nonlinear minimization problem to determine the absorption coefficient μ_a in V_{N_e}:

$$min_{\mu_a \in V_{N_e}} F(\mu_a), \tag{2.7}$$

$$F(\mu_a) = \frac{1}{2}\sum_{j=1}^{N_s}\sum_{i=1}^{N_d} (\Gamma_{i,j}(\mu_a) - \mathcal{M}_{i,j})^2 + \frac{\alpha}{2}\sum_{m=1}^{N_e} \mu_a(m)^2,$$

where $\mathcal{M}_{i,j}$ are measured data at the i-th detector and the j-th source, and α is a regularization parameter.

Although the uniqueness of μ_a (and μ_s') is known for infinite sources and detectors, the unique solvability of (2.7) for finite sources and detectors is not known leading to the ill-posedness of the DOT problem. To address this ill-posedness, nonlinear optimization such as Newton method is employed.

The optimization problem in (2.7) is composed of two steps; the step to determine the minimizing direction at the current coefficients μ_a and the step to perform line search on those minimizing directions. In the Newtonian method, the minimizing direction is $-(F'')^{-1}F'$, where F' and F'' are the Jacobian and the Hessian of F with respect to μ_a. In the classical Newton method, searching direction is found by computing F'' directly. However, this method requires extensive computation and is unstable. To decrease complexity of the computation, in quasi-Newton methods, F''^{-1} is approximated by various computationally efficient methods such as the Davidon-Fletcher-Powell method, the Broyden-Fletcher-Goldfarb-Shanno method, and the conjugate gradient method [7]. To stabilize the computation, $F''+\lambda I$ is used in place of F'' in the Levenberg-Marquadt method [16, 17].

In this paper, we adopted the trust region method, a variant of Newton approach, as an optimizer. The trust region method solves the following quadratic minimization problem:

$$min\left\{\delta\mu_a \in W, \|D\delta\mu_a\| \leq T \left| \frac{1}{2}(\delta\mu_a)^t F''(\mu_a)\delta\mu_a + (\delta\mu_a)^t F'(\mu_a)\right.\right\} \qquad (2.8)$$

where D is a scaling matrix, T is a trust region parameter, and W is a subspace of V. The scaling matrix D is used to handle constraints for the minimization. To stabilize the minimization, we control the trust region parameter T, which is similar to control the λ parameter in the Levenberg-Marquadt method [7]. To avoid extensive computation, the subspace W is chosen as the two-dimensional subspace composed of the gradient direction and the approximate Newton direction [9]. This makes the trust region method suitable for large-scale constrained optimization problems like the one state in (2.7).

Taking vanishing gradient point of the quadratic form $\frac{1}{2}(\delta\mu_a)^t F''(\mu_a)\delta\mu_a + (\delta\mu_a)^t F'(\delta\mu_a)$ in (2.8), we get

$$\left(J^t(\delta\mu_a)J(\delta\mu_a) + \alpha I_{N_e}\right)\delta\mu_a = -(J(\mu_a)^t b(\mu_a) + \alpha\mu_a). \qquad (2.9)$$

Thus, if assume that $\delta\mu_a$ is sufficiently small, the trust region method can be used to solve (2.8) at each iteration.

3. Two-level Domain decomposition methods for the diffuse optical tomography

In this section, we describe the two-level domain decomposition methods considered in this paper, as applied to the forward and inverse problems. These two methods are composed of the coarse level and the subdomain correction steps. We will describe our notation and approach for the two-dimensional optical domain $\Omega = [a,b] \times [c,d] \subset \mathbb{R}^2$ and bilinear finite element. Its extension to three-dimensional domain is straightforward.

Let Ω_{N_x,N_y} be Ω uniformly divided by N_x times in the x-axis direction and N_y times in the y-axis direction. Thus Ω_{N_x,N_y} has $N_n = (N_x+1)\times(N_y+1)$ nodes and $N_e = N_x \times N_y$ elements. We call $\Omega_{N_x/2,N_y/2}$ the coarse level of Ω_{N_x,N_y} assuming N_x and N_y are even. Let Ω_{N_x,N_y} be decomposed into disjoint union of d subdomains $\Omega_p, p = 1,\cdots,d$ such that

$$\Omega_{N_x,N_y} = \bigcup_{p=1}^{d} \overline{\Omega}_p. \tag{3.10}$$

(3.10) describes the nonoverlapping domain decomposition. For the overlapping domain decomposition, we define $\Omega_p^w, p = 1,\cdots,d$, an extension of Ω_p, recursively for all nonnegative integers w as follows: $\Omega_p^0 = \Omega_p$ and Ω_p^w is the union of Ω_p^{w-1} and its adjacent elements in contact with the boundary of Ω_p^{w-1}, where w will be called the overlapping width. Thus the overlapping domain decomposition is given by

$$\Omega_{N_x,N_y} = \bigcup_{p=1}^{d} \Omega_p^w. \tag{3.11}$$

3.1. Two-level Multiplicative Overlapping Domain Decomposition Method for the forward problem

In the forward solution of the DOT problem, one has to compute the photon densities satisfying (2.1) and its Jacobian. Commonly used methods are the finite element method in (2.5) and the adjoint method (2.6). We apply TMODDM to (2.5) in order to reduce the computation complexity of the problem. TMODDM composed of two steps: the coarse level correction and the subdomain correction steps.

In the subdomain correction step, we use MODDM. Let Φ^n be the n-th step solution of TMODDM and $\Phi^{n+\frac{1}{d+1}}$ be the updates via coarse level correction to be presented below. Then the updates $\Phi^{n+\frac{p+1}{d+1}}$ at p-th subdomain Ω_p^w is obtained by

$$\Phi^{n+\frac{p+1}{d+1}} = \begin{cases} v & \text{in } \Omega_p^w \\ \Phi^{n+\frac{p}{d+1}} & \text{in } \Omega \setminus \overline{\Omega}_p^w \end{cases} \tag{3.12}$$

where v is the solution of the following partial differential equation:

$$-\nabla \cdot (\kappa \nabla v) + (\mu_a + \frac{i\omega}{c})v == q \qquad \text{in } \Omega_p^w \tag{3.13a}$$

$$v + 2a\kappa\frac{\partial v}{\partial \nu} = 0 \qquad \text{on } \partial\Omega_p^w \cap \partial\Omega, \tag{3.13b}$$

$$v = \Phi^{n+\frac{p}{d+1}} \qquad \text{on } \partial\Omega_p^w \setminus \partial\Omega. \tag{3.13c}$$

We call the coarse level correction step combined with MODDM TMODDM. In the coarse level correction step, an important point is to define the interpolation operator

$P : \Omega_{N_x/2,N_y/2} \rightarrow \Omega_{N_x,N_y}$ and restriction operator $R : \Omega_{N_x,N_y} \rightarrow \Omega_{N_x/2,N_y/2}$ between the functions defined at the node points in $\Omega_{N_x/2,N_y/2}$ and Ω_{N_x,N_y}.

We use bilinear interpolation operator P and full weighting restriction operator R, having the following symbols:

$$P = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}, \text{ and } R = \frac{1}{4}P^T,$$

where P^T is the transpose of P[8, 24]. At the boundary, P and R has the symbol $(1/2 \quad 1 \quad 1/2)$ boundary is a one-dimensional line.

Let the maximum sweep of all subdomain corrections followed by the coarse grid correction be M_F. A pseudo-code summarizing the Algorithm TMODDM to solve (2.5) at Ω_{N_x,N_y} is given as follows:

Algorithm 1 TMODDM

```
for $j = 1 \cdots N_s + N_d$ do
  Initialize $\Phi_j^0$.
  for $n = 1 \cdots M_F$ do
    $r^j \Leftarrow R((K + C + \frac{1}{2a}A)\Phi_j^{n-1} - b^j)$, $r^j \in \Omega_{N_x/2,N_y/2}$ {Restrict the residual.}
    $e_j \Leftarrow (K + C + \frac{1}{2a}A)^{-1} r^j$, $e_j, r^j \in \Omega_{N_x/2,N_y/2}$ {Coarse grid correction.}
    $\Phi_j^{n-1+\frac{1}{d+1}} \Leftarrow \Phi_j^{n-1} + P(e_j)$, $P(e_j) \in \Omega_{N_x,N_y}$
    for $p = 1, \cdots, d$ do
      Update $\Phi_j^{n-1+\frac{p+1}{d+1}}$ at $\Omega_p^w$ by (3.12) {Subdomain correction.}
    end for
  end for
end for
Compute $\Gamma_{i,j}$ and $\frac{\partial \Gamma_{i,j}}{\partial \mu_a}(m), m = 1, \cdots, N_e$ using (2.4) and (2.6). {Post-processing}
```

3.2. Two-level multiplicative space decomposition method for the inverse problem

In the forward problem, we obtain the value of photon density at each nodes. Whereas in the inverse problem, we are interested in the value of the optical coefficients at each element. It is possible to define different domain decomposition schemes for the nodes and elements in the forward and inverse problems. In this paper, however, we use the same domain decomposition scheme for both problems. For the domain decomposition (3.10) and (3.11), let the restrictions of V_{N_e} to the overlapping subdomain Ω_p^w and nonoverlapping subdomain Ω_p be $V_{N_e}^p$ and $W_{N_e}^p$ for $p = 1, \cdots, d$, respectively, such that

$$V_{N_e}^p = \{x \in V_{N_e} | x = 0 \text{ on } \Omega \setminus \overline{\Omega_p^w}\}, \tag{3.14a}$$

$$W_{N_e}^p \quad = \quad \{x \in W_{N_e} | x = 0 \text{ on } \Omega \setminus \overline{\Omega_p}\}. \tag{3.14b}$$

TMSDM composed of two steps: The coarse level correction and the subdomain correction steps.

The subdomain correction step uses the algorithm MSDM which is given below: Let μ_a^0 be the initial guess for optical coefficient on Ω_{N_x,N_y}. Define $\mu_a^{n+p/d}$ as the n-th update of the optical coefficient in the p-th subdomain. Then given $\mu_a^{n+\frac{p-1}{d}} \in V_{N_e}^p, p = 1, \cdots, d$, compute $\delta\mu_a^{n+\frac{p}{d}}$ such that

$$\delta\mu_a^{n+\frac{p}{d}} = argmin_{\delta\mu_a^{n+\frac{p}{d}} \in V_{N_e}^p, \mu_a^{n+\frac{p-1}{d}} + \delta\mu_a^{n+\frac{p}{d}} \in V_{N_e}} F(\mu_a^{n+\frac{p-1}{d}} + \delta\mu_a^{n+\frac{p}{d}}), \tag{3.15}$$

and update

$$\mu_a^{n+\frac{p}{d}} = \mu_a^{n+\frac{p-1}{d}} + \delta\mu_a^{n+\frac{p}{d}}. \tag{3.16}$$

The convergence of MSDM is proved under some conditions on the objective function F [5]. We have proved the convergence of MSDM in Theorem 1 when approximating the μ_a in each iteration is sufficiently close to the true μ_a. As a result, the convergence of the MSDM is assured when the initial guess is sufficiently close to the true μ_a. Motivated by the Algorithm TMODDM, we applied a two-level approach to the MSDM by using the coarse level solution as an initial guess. For the present work, we assume that the coarse level solution is sufficiently close to the true image of the absorption coefficient.

To use the coarse level solution in the finer level as an initial guess, we adopted an upsampling operator U from the coarse level to the fine level. Let y be a function defined on $N_x/2 \times N_y/2$ elements on the coarse level, then $U(y)$ is a function defined on each $N_x \times N_y$ elements on the fine level such that

$$\begin{aligned} U(y)(2m_x - 1, 2m_y - 1) &= y(m_x, m_y), \quad & U(y)(2m_x - 1, 2m_y) &= y(m_x, m_y), \\ U(y)(2m_x, 2m_y - 1) &= y(m_x, m_y), \quad & U(y)(2m_x, 2m_y) &= y(m_x, m_y), \end{aligned}$$

for $1 \leq m_x \leq N_x/2, 1 \leq m_y \leq N_y/2$.

Let M_C and M_S denote the number of sweeps for the coarse and subdomain corrections. A pseudo-code summarizing the Algorithm TMSDM is given as follows:

Assume that we know the location of the anomalous region is contained in some small region by a priori information provided by a secondary imaging modality such as MRI or X-rays or by a posteriori information after the coarse level correction step. Then instead of performing the subdomain correction step on all of the subdomains, we can update μ_a on a region of anomaly. The convergence of the algorithm when implemented only for a region of interest does not change if the initial guess of the optical coefficient is sufficiently close to the true image of the coefficient.

Algorithm 2 TMSDM

```
$\Gamma^c = \log(R(exp(\Gamma)))$ {Restrict boundary measurements.}
for $i = 1, \cdots, M_C$ do
  Solve (2.8) on the coarse level from the boundary measurement data $\Gamma^c_{i,j}$ to
  approximate $\mu^c_a$ {Coarse level correction.}
end for
$\mu_a \Leftarrow U(\mu^c_a)$
for $p = 1, \cdots, d$ do
  for $k = 1, \cdots, M_S$ do
    Solve (3.15) from the measurement data $\Gamma_{i,j}$ and elaborate $\mu_a$ on $\Omega^w_p$ by (3.16)
    using Algorithm 1 as a forward solver {Subdomain correction.}
  end for
end for
```

3.3. Convergence of TMODDM and TMSDM

In this subsection, we will discuss the convergence behavior of the proposed algorithms. Assume that the mesh size of the finite element formulation is $O(h)$ and that the subdomains are of diameter $O(H)$ and the overlap region is width $O(\delta H)$. Then the following convergence behavior is known for the Algorithm TMODDM [20, 19].

- Convergence is poor if $\delta = 0$ but improves rapidly as δ increase. (3.17a)
- If δ is fixed, the number of iterations is bounded independent of h, H and H/h. (3.17b)
- The number of iterations for the multiplicative Schwarz method is roughly half of that needed for the additive Schwarz method. (3.17c)

The local linear convergence of the Algorithm MSDM is shown below using the results in [5].

Theorem 1 Let μ^n_a are the n-th step MSDM approximation of μ_a. Assume that $\|\mu_a - \mu^q_a\|_{V_{N_e}} \le \delta(q = 0, \cdots, n)$ for sufficiently small δ. Then, we have

$$\|\mu^n_a - \mu_a\|_{V_{N_e}} \le C_1(C_2)^n |F(\mu^0_a) - F(\mu_a)|, \tag{3.18}$$

where C_1 is a positive constant and C_2 is a positive constant less than 1. (See appendix for the proof of Theorem 1).

3.4. Computational cost of the Algorithms

In this subsection, the computational cost of the proposed Algorithms will be estimated when solving (2.5) and (2.9) for the forward and inverse problem, respectively. We will call, the method solving (2.5) and (2.9) without any domain decomposition, non-DDM compared to the proposed method. The system matrices for each equation are $N_n \times N_n$ and $N_e \times N_e$, respectively. Assume that we require $O(N_n^p)$ and $O(N_e^p)$ floating point operations for each equation with non-DDM. For full nonzero matrix we must take $p = 3$ and for usual non sparse matrix we can take $2 \leq p \leq 3$. Suppose that we use d subdomains with equal nodes and elements for the proposed algorithms.

The coarse level computation needs 4^{-p} times (in two-dimension) or 8^{-p} times (in three-dimension) less computation than the fine level computation for the forward and inverse solvers. Thus if we neglect the coarse level computation, the computational cost of the Algorithm MODDM is $M_F d^{1-p}$ (for one computer) and $M_F d^{-p}$ (for d parallel computers) times that of non-DDM method, for one computer and for d parallel computers, respectively. It is well known that the number of iterations M_F is bounded independent of N_n, d and N_n/d by (3.17) and TMODDM gives a good convergence behavior with smaller M_F than MODDM. Thus we have chosen $M_F = 3$ when TMODDM is used in the inverse solver in Section 4.2. Thus, by using multiple subdomains, the Algorithm TMODDM can achieve significant decrease in computational requirements.

The total computation costs for non-DDM, the Algorithm TMSDM with forward solver the Algorithm TMODDM in case of one computer, d parallel computers, and d^2 parallel computers are tabulated in Table 1. When we use d parallel computers, the parallel computing is applied to the inverse solver only, whereas when we use d^2 computers, the parallel computing is applied to the inverse and forward solvers simultaneously. Note that this comparison for the inverse solver is based on the assumption that the speed of data communications between parallel computers is sufficiently fast. The parallel computation is not treated in this paper, but the proposed algorithm combined with the parallel computing will lead to much more efficient results.

4. Numerical Simulations

In this section, we will test the efficiency of the two algorithms: Algorithm TMODDM for the forward problem and Algorithm TMSDM for the inverse problem using simulated data.

The following are the parameter values chosen for the numerical simulations.

$$\begin{aligned} \mu_s' &= 8cm^{-1}, & \text{background } \mu_a &= 0.05cm^{-1}, \\ \text{tumor } \mu_a &= 0.2cm^{-1}, & \omega &= 2\pi * 100MHz, \\ c &= 3 * 10^{10} cm/sec, & a &= 1. \end{aligned}$$

In Figure 1(a), the thin lines represent 20×20 fine level and thick lines represent 10×10 coarse level for $\Omega = [0,6] \times [0,6]$. The 2×2 domain decomposition and 4×4

domain decomposition with $w = 1$ are shown in Figure 1 (b) and (c). 3×1 domain decomposition for $\Omega = [0,9,6] \times [0,4.8]$ with 24×18 node points is shown in Figure 1 (d).

4.1. Algorithm TMODDM

In this subsection, we consider Algorithm TMODDM in Figure 1, Table 2, and Table 3. Let $\Omega = [0,6] \times [0,6]$ be divided into 20×20 square elements. The distribution of μ_a is shown in Figure 2(a), where the white region is the background with $\mu_a = 0.05cm^{-1}$ and the black region is the anomaly with absorption coefficient $\mu_a = 0.2cm^{-1}$. Let the location of the source be $(6,3)$. The solution of (2.5) by LINPACK is in Figure 2(b). The $L^2(\Omega)$ errors for the photon density between LINPACK solution and the solution of the proposed method, are displayed in Figures 2(c)-(e), Table 2, and 3. In Table 2, the error for one-level and two-level MODDM is compared as the number of iterations increases for 2×2 subdomains as in Figure 1(b). Although TMODDM 4^{-p} more computation than MODDM, the convergence behavior of the proposed method for each iteration is remarkably good as compared to the one-level method. In Table 3, the effect of the number of subdomains and the width of overlapping region is tabulated. As the width gets larger, the error decreases with a few exceptions. This verifies the convergence result of (3.17a). In Figure 2(c),(d), and (e), the L^2 error is shown after 5 iterations, after the coarse level correction after 5 iterations, and after 6 iterations, respectively. The errors for each step are 2.51×10^{-9}, 8.95×10^{-10}, and 8.31×10^{-11}, respectively.

4.2. Algorithm TMSDM

In this subsection, we evaluate the performance of the algorithm TMSDM using simulated data and Algorithm TMODDM as a forward solver. The results are shown in Figures 3-5. The reconstruction of diffusion coefficient κ for known μ_a is investigated in [13]. In this paper, the reconstruction of μ_a is implemented when μ_s' is given as in [14].

Algorithm TMSDM is implemented for various parameters including type of domain Ω, μ_a, source and detector locations, and the Tikhonov regularization parameter. As a forward solver, we have used the Algorithm TMODDM with maximum number of iterations set to 3. And if we know the approximate location of tumor a priori, we only updated μ_a only on some small region of interest instead of all subdomains of Ω in inverse solver.

The maximum number of iterations for the coarse level correction, the subdomain correction, and non-DDM is 10, 5, and 10, respectively. As shown in Table 1, the computational cost of the proposed method is less than that of non-DDM. Figures ??(a), 4(a), and 5(a), show the domain Ω and μ_a images used in simulations. Figures

3(b), 4(b), and 5(b), the location of the sources and detectors. 'o' and '*' mark the location of sources and detectors, respectively. Figures 3(c)-(e), 4(c)-(f), and 5(c)-(e), show the reconstructed images of μ_a for various parameters and the $L^2(\Omega)$ norm error between the original μ_a and reconstructed μ_a.

In Figure 3, we considered $\Omega = [0,6] \times [0,6]$ domain divided into 400 pixels with $N_x = N_y = 20$. We choose 2×2 subdomains of the same size with overlapping width 1 as shown in Figure 1(b). Figure 3(a) shows the original image of μ_a. 20 sources are interweaved with 20 detectors located on the boundary as shown in Figure 3(b). The image is reconstructed using the Algorithm TMSDM. In Figures 3(c) and 3(d), the reconstructed images are shown for Tikhonov regularization parameter $\alpha = 0$ and 10^{-2}, respectively. As compared to the image with no regularization, the image with $\alpha = 10^{-2}$ appears sharper and the $L^2(\Omega)$ norm error is decreased from 0.145 to 0.129. The reconstructed image for the non-DDM after 10 iterations is shown in Figure 3(e). The boundary of the anomaly does not appear clearly and the image of anomaly is larger than the original image, and the $L^2(\Omega)$ error is larger than that of the DDM approach shown in Figure 3(d).

In Figure 4, the same domain Ω, pixels, subdomains, and source and detector locations are considered as in Figure 3. But the anomaly of μ_a is located at the upper right quadrant of Ω as shown in Figure 4(a). The reconstructed images using Algorithm TMSDM with $\alpha = 0$ and $\alpha = 10^{-2}$ are shown in Figure 4(c) and Figure 4(d). The reconstructed image when μ_a is updated only in 4-th subdomain is shown in Figure 4(e). The reconstructed image for the non-DDM is shown in Figure 4(f). The $L^2(\Omega)$ error becomes smaller when we use $\alpha = 10^{-2}$ in Figures 4 (d)-(f) compared to $\alpha = 0$ in Figure 4 (c). As compared to Figure 4 (d), when we only update the 4-th subdomain, the error in Figure 4 (e) decreases significantly. Thus if a priori or a posteriori information about the location of tumor is available, it is possible to obtain more accurate and faster reconstruction as shown in Figure 4(e) rather than Figure 4(d). The reconstructed image for the non-DDM in Figure 4(f) is somewhat blurred and the boundary is not so clear as compared to the images obtained using Algorithm TMSDM.

In Figure 5, we considered a thin slab geometry with $\Omega = [0,9.6] \times [0,4.8]$ divided into 288 pixels with $N_x = 24$ and $N_y = 12$. We have choose 3×1 subdomains of similar size with overlapping width as shown in Figure 1(d). 13 sources on the bottom and 13 detectors on the roof of Ω are located as shown in Figure 5(b). The reconstructed images with $\alpha = 0$ and $\alpha = 10^{-2}$ are shown in Figure 5(c) and (d). Figure 5(e) is the reconstructed image with non-DDM. The reconstructed image with regularization shown in Figure 5(d) has smaller $L^2(\Omega)$ error than the one without any regularization shown in Figure 5(e). As in Figure 3 and 4, the reconstructed image for the non-DDM shows blurred boundary of the anomaly.

5. Conclusion and discussion

We presented Algorithm TMODDM and Algorithm TMSDM for the forward and the inverse problem solution of DOT. The algorithms of these two methods are explained and their convergences are shown. Numerical implementations of the algorithms are presented for various parameters such as geometry, source-detector location, Tikhonov parameter, the number and width of subdomain. For Algorithm TMODDM, the forward solver, as overlapping width grows the convergence becomes better for the 4 and 16 subdomains, and the iteration number for the proposed method decreases fundamentally compared to the one-level method. For Algorithm TMSDM combined with Algorithm TMODDM, Tikhonov parameter 10^{-2} produces less $L^2(\Omega)$ error for the absorption coefficient than other Tikhonov parameters. The images of the coefficient by non-DDM are blurred and makes higher $L^2(\Omega)$ error than the proposed method. Using Algorithm TMSDM in a region of interest produces better quality images and less error as compared to the Algorithm TMSDM applied in all subdomains.

The computational costs for the two-level methods are compared among non-DDM, the proposed method in 1, d, and d^2 parallel computers, where d is the number of subdomains. Thus, with the aid of parallel computing for d^2 computers, the proposed approach speeds up the computation with d^{-p} times as compared to the non-DDM, where usually p is between 2 and 3 depending on the system matrix size and matrix solver. If a region of interest is determined in some subdomain by a priori information obtained from secondary imaging modality such as X-ray, CT, MRI or a posteriori information from a coarse level image, Algorithm TMSDM can be implemented in the region of interest resulting in a d^{-p} speed-up of computation only in d computers.

6. Appendix: the proof of Theorem 1

More detailed proof of the Theorem in general case including the concrete definitions of Jacobian and Hessian, and their error estimates, will be studied in [15]. In this appendix, we use absolute intensity $\Gamma = \frac{1}{2a}\Phi$ as a measurement data for real valued Φ instead of log intensity in (2.2), and assume that $\kappa = \frac{1}{3\mu_s'}$, which means $\frac{\partial\kappa}{\partial\mu_a} = 0$.

For i-th source r_i, j-th detector r_j and elements T_k and T_l, let $b_{ij} = \Gamma_{ij}(\mu_a + \delta\mu_a) - \Gamma_{ij}(\mu_a)$. The Jacobian matrix J and Hessian tensor H at a given μ_a are given by

$$J(\mu_a)[\delta\mu_a](ij,k) = \int_{T_k} G(r_i,r')G(r',r_j)\delta\mu_a(r')dr'$$

$$H(\mu_a)[\delta\mu_a,\delta\mu_a](ij,k,l)$$

$$= \int_{T_k}\int_{T_l} G(r_i,r')G(r',r'')G(r'',r_j)\delta\mu_a(r')\delta\mu_a(r'')dr'dr''$$

where $G(\cdot, r_j)$ is the Green function for μ_a and a point source r_j. Define

$$\|J\| := \max_{i,j,k} |J(ij,k)| \le \sup_{r\in\overline{\Omega}} \|G(\cdot,r)\|^2_{L^2(\Omega)} \|\delta\mu_a\|_{L^\infty(\Omega)},$$
$$\|H\| := \max_{i,j,k,l} |H(ij,k,l)| \le \sup_{r\in\overline{\Omega}} \|G(\cdot,r)\|^3_{L^2(\Omega)} \|\delta\mu_a\|^2_{L^\infty(\Omega)}.$$

Lemma 2 Assume that $\delta\mu_a$ is small enough to satisfying $\|H\|\,\|J\|\,\|\delta\mu_a\|_{L^\infty(\Omega)} \le \frac{\alpha}{2}$ and $\sup_{r\in\overline{\Omega}} \left\|\widetilde{G}(\cdot,r)\right\|_{L^2(\Omega)} \le \frac{3}{2}\sup_{r\in\overline{\Omega}} \|G(\cdot,r)\|_{L^2(\Omega)}$, where $\widetilde{G}(\cdot,r)$ is the solution of (2.1) for the absorption coefficient $\mu_a + \delta\mu_a$. Then F'' satisfies

$$\frac{\alpha}{2}\|\delta\mu_a\|^2_{V_{N_e}} \le (\delta\mu_a)^t F''(x)(\delta\mu_a). \tag{6.19}$$

Proof F' and F'' are computed as follows:

$$F'(\mu_a) = J(\mu_a)^t b(\mu_a) + \alpha\mu_a,$$
$$F''(\mu_a) = J(\mu_a)^t J(\mu_a) + H(\mu_a) b(\mu_a) + \alpha I.$$

By Born expansion

$$|b_{ij}| = |\widetilde{G}(r_{i+N_s}, r_j) - G(r_{i+N_s}, r_j)| = |\int_\Omega G(r_i,r')\widetilde{G}(r',r_j)\delta\mu_a|,$$

where $\widetilde{G}$ is the solution of (2.1) for $\mu_a + \delta\mu_a$. Using the assumption and the definition of J, we get $|b_{ij}| \le \frac{3}{2}\|J\|\,\|\delta\mu_a\|_{L^\infty(\Omega)}$. Therefore we get

$$\begin{aligned}\frac{\alpha}{2}\|\delta\mu_a\|_{V_{N_e}} &\le \left(\alpha - \frac{3}{2}\|H(\mu_a)\|\,\|J(\mu_a)\|\,\|\delta\mu_a\|_{L^\infty(\Omega)}\right)\|\delta\mu_a\|^2_{V_{N_e}} \\ &\le (\delta\mu_a)^t F''(\mu_a)\delta\mu_a.\end{aligned} \tag{6.20}$$

Thus we have proved the lemma.

The inequality (6.19) is called local strong convexity. With this property and the theorem in [5], we prove Theorem 1 for δ is sufficiently small such that $\delta\mu_a = \mu_a^q - mu_a, q = 0,\cdots,n$ satisfies the assumption of Lemma 2.

Proof of Theorem 1 The proof of Theorem 1 is the modification of the proof in [5]. By (3.14),

$$V_{N_e} = V^1_{N_e} + \cdots + V^d_{N_e} = W^1_{N_e} + \cdots + W^d_{N_e}$$

and $W^p_{N_e}, p = 1,\cdots,d$ are mutually disjoint and $W^p_{N_e} \subset V^p_{N_e}$. And let and ${\mu_a}^{n+k/d}$ be the minimizing solution of TMSDM at n-th step and p-th subdomain Ω^w_p. Define $z^n_p = (\mu_a - {\mu_a}^{n+\frac{p}{d}})|_{\Omega_p} \in W^p_{N_e}$ and $e^{n+p/d} = {\mu_a}^{n+\frac{p}{d}} - {\mu_a}^{n+\frac{p-1}{d}} \in V^p_{N_e}$. Then we get

$$\mu_a - {\mu_a}^n = \sum_{p=1}^d z^n_p, \quad \|\mu_a - {\mu_a}^n\|_{V_{N_e}} = \left(\sum_{p=1}^d \|z^n_p\|^2_{V_{N_e}}\right)^{1/2} \tag{6.21}$$

and

$$e^{n+\frac{p}{d}} = \arg\, \min_{v_p \in {V_{N_e}}^p} F({\mu_a}^{n+\frac{p-1}{d}} + v_p). \tag{6.22}$$

(6.22) implies

$$\left\langle F'({\mu_a}^{n+\frac{p-1}{d}} + e^{n+\frac{p}{d}}), v_p - e^{n+\frac{p}{d}} \right\rangle \geq 0 \text{ for all } v_p \in {V_{N_e}}^p. \tag{6.23}$$

Using (6.19) and Taylor expansion for F we get

$$F(w) - F(v) \geq \langle F'(v), w - v \rangle + \frac{\alpha}{4} \|w - v\|_{V_{N_e}}^2 \tag{6.24}$$

for all $v, w \in V_{N_e}$ such that $\|v - w\|_{V_{N_e}} \leq \delta$. Taking $v_p = 0$ in (6.23) and inserting $w = {\mu_a}^{n+\frac{p-1}{d}}$ and $v = {\mu_a}^{n+\frac{p}{d}}$ we get

$$F({\mu_a}^{n+\frac{p-1}{d}}) - F({\mu_a}^{n+\frac{p}{d}}) \geq \frac{\alpha}{4} \left\| e^{n+\frac{p}{d}} \right\|_{{V_{N_e}}^p} \geq 0$$

and

$$\begin{aligned} F({\mu_a}^n) - F({\mu_a}^{n+1}) &\geq \sum_{p=1}^{d} \left(F({\mu_a}^{n+\frac{p-1}{d}}) - F({\mu_a}^{n+\frac{p}{d}}) \right) \\ &\geq \frac{\alpha}{4} \sum_{p=1}^{d} \left\| e^{n+\frac{p}{d}} \right\|_{{V_{N_e}}^p}^2 \geq 0. \end{aligned} \tag{6.25}$$

Inserting $w = {\mu_a}^n$ and $v = \mu_a$ into (6.24), we get

$$\begin{aligned} F({\mu_a}^n) - F(\mu_a) &= \langle F'(\mu_a), {\mu_a}^n - \mu_a \rangle + \frac{\alpha}{4} \|{\mu_a}^n - \mu_a\|_{V_{N_e}}^2 \\ &\geq \frac{\alpha}{4} \|{\mu_a}^n - \mu_a\|_{V_{N_e}}^2 \geq 0. \end{aligned} \tag{6.26}$$

Finally we get

$$\begin{aligned} &0 \leq F({\mu_a}^{n+1}) - F(\mu_a) \leq \left\langle F'({\mu_a}^{n+1}), {\mu_a}^{n+1} - \mu_a \right\rangle \text{ (using (6.24))} \\ &= \left\langle F'({\mu_a}^{n+1}), \sum_{p=1}^{d} e^{n+\frac{p}{d}} + {\mu_a}^n - \mu_a \right\rangle = \sum_{p=1}^{d} \left\langle F'({\mu_a}^{n+1}, e^{n+\frac{p}{d}} - z_p^n \right\rangle \\ &\leq \sum_{p=1}^{d} \left\langle F'({\mu_a}^{n+1}) - F'({\mu_a}^{n+\frac{p}{d}}), e^{n+\frac{p}{d}} - z_p^n \right\rangle \text{ (using (6.23))} \\ &= \sum_{p=1}^{d} \sum_{l=p+1}^{d} \left\langle F'({\mu_a}^{n+\frac{l}{d}}) - F'({\mu_a}^{n+\frac{l-1}{d}}), e^{n+\frac{p}{d}} - z_p^n \right\rangle \\ &= \sum_{p=1}^{d} \sum_{l=p+1}^{d} \left\langle F''(y_l^n) e^{n+\frac{l}{d}}, e^{n+\frac{p}{d}} - z_p^n \right\rangle \end{aligned}$$

for some y_l^n which lies between ${\mu_a}^{n+\frac{l}{d}}$ and ${\mu_a}^{n+\frac{l-1}{d}}$

$$\leq \frac{\alpha}{2}\left(\sum_{l=1}^{d}\left\|e^{n+\frac{l}{d}}\right\|^2_{V_{Ne}}\right)^{1/2}\left(\sum_{p=1}^{d}\left\|e^{n+\frac{p}{d}}-z_p^n\right\|^2_{V_{Ne}}\right)^{1/2} \quad \text{(using (6.19))}$$

$$\leq \frac{\alpha}{2}\left[\sum_{l=1}^{d}\left\|e^{n+\frac{l}{d}}\right\|^2_{V_{Ne}}+\left(\sum_{p=1}^{d}\left\|e^{n+\frac{p}{d}}\right\|^2_{V_{Ne}}\right)^{1/2}\left(\sum_{p=1}^{d}\left\|z_p^n\right\|^2_{V_{Ne}}\right)^{1/2}\right]$$

$$\leq 2(F({\mu_a}^n)-F({\mu_a}^{n+1}))+2\sqrt{F({\mu_a}^n)-F({\mu_a}^{n+1})}\sqrt{F({\mu_a}^n)-F(\mu_a)}$$

(using (6.21), (6.25), and (6.26)).

Let $d_n = F({\mu_a}^n) - F(\mu_a)$, then the above equation and Cauchy-Schwarz inequality implies

$$\begin{aligned} d_{n+1} &\leq 2(d_n - d_{n+1}) + 2\sqrt{d_n - d_{n+1}}\sqrt{d_n} \\ &\leq (2 + \frac{1}{2\mu})(d_n - d_{n+1}) + 2\mu d_n, \end{aligned}$$

for all $\mu > 0$. Thus we get

$$\frac{d_{n+1}}{d_n} \leq \frac{4\mu^2 + 4\mu + 1}{6\mu + 1} \tag{6.27}$$

The left hand side of (6.27) takes minimum value $\frac{8}{9} < 1$ when $\mu = \frac{1}{6}$. Thus if we take $C_1 = \frac{4}{\alpha}$ and $C_2 = \frac{8}{9}$, Theorem 1 is proved.

7. Acknowledgement

This material is based upon work supported by NSF-BES-0353160, ONR-N00014-04-1-0694, and US Army Medical Research-W81XWH-04-1-0559. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Various portions of this research were supported by the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821), and Rensselaer Polytechnic Institute.

[1] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–93, 1999.

[2] S. R. Arridge. Methods for the inverse problem in optical tomography. In P. Sebbah and J.-M. Tualle, editors, *Waves and Imaging Through Complex Media*, pages 307–329. Kluwer, 2001.

[3] S. R. Arridge, M. Schweiger, and D. T. Delpy. Iterative reconstruction of near-infrared absorption images. In M. A. Fiddy, editor, *Inverse problems in Scattering and Imaging*, volume 1767 of *Proceedings of SPIE*, pages 372–383, 1992.

[4] L. Badea. On the Schwarz alternation method with more than two subdomains for nonlinear monotone problems. *SIAM J. Numer. Anal.*, 28:179–204, 1991.

[5] L. Badea, X.-C. Tai, and J. Wang. Convergence rate analysis of a multiplicative Schwarz method for variational inequalities. *SIAM J. Numer. Anal.*, 41:1052–73, 2003.

[6] L. Badea and J. Wang. An additive Schwarz method for variational inequalities. *Mathematics of Computation*, 69:1341–54, 1999.

[7] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming, theory and algorithms.* John Wiley & Sons, Inc., second edition, 1993.

[8] W. L. Briggs. *A Multigrid Tutorial.* SIAM, 1987.

[9] R. H. Byrd, R. B. Schnabel, and G. A. Shultz. Approximate solution of the trust region problem by minimization over two-dimensional subspaces. *Mathematical programming*, 40:247–63, 1988.

[10] V. Isakov. On uniqueness in the inverse transmission scattering problem. *Commun. Partial Diff. Eq.*, 15:1565–87, 1990.

[11] K. Kwon. Identification of anisotropic anomalous region in inverse problems. *Inverse Problems*, 20:1117–36, 2004.

[12] K. Kwon and D. Sheen. Anisotropic inverse conductivity and scattering problems. *Inverse Problems*, 18:745–56, 2002.

[13] K. Kwon, I.-Y. Son, and B. Yazici. Domain decomposition method for diffuse optical tomography. In Charles A. Bouman and Eric L. Miller, editors, *Computational Imaging III*, volume 5674 of *Proceedings of SPIE*, pages 64–75, 2005.

[14] K. Kwon, I.-Y. Son, and B. Yazici. Two-level overlapping domain decomposition algorithm for a nonlinear inverse DOT problem. In B. Chance, R. R. Alfano, B. J. Tromberg, M. Tamura, and E. M. Sevick-Muraca, editors, *Optical Tomography and Spectroscopy of Tissue VII*, volume 5693 of *Proceedings of SPIE*, pages 459–468, 2005.

[15] K. Kwon and B. Yazici. Higher derivatives of measurement vector in diffuse optical tomography. in preparation.

[16] K. Levenberg. A method of the solution of certain problems in least squares. *Quart. Appl. Math.*, 2:164–68, 1944.

[17] D. Marquadt. An algorithm for least-squares estimation of nonlinear parameters. *IEEE Transactions of Circuits and Systems*, CAS-26,Sept., 1979.

[18] M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh. Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography. *Opt. Lett.*, 20:426–428, 1995.

[19] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations.* Oxford Science Publications, 1999.

[20] B. Smith, P. Bjørstad, and W. Gropp. *Domain Decomposition, Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, 1996.

[21] I.-Y. Son, M. Guven, B. Yazici, and X. Intes. A 2-level domain decomposition algorithm for inverse diffuse optical tomography. In *Proceedings of ICIP.* IEEE, 2004.

[22] J. Sylvester and G. Uhlmann. A global uniqueness theorem for an inverse boundary value problem. *Ann. math.*, 125:153–69, 1987.

[23] X.-C. Tai. Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities. *Numer. Math.*, 93:755–86, 2003.

[24] P. Wesseling. *An introduction to multigrid methods.* John Wiley & Sons, 1992.

[25] J.-P. Zeng and S. Z. Zhou. A domain decomposition method for a kind of optimization problems. *Journal of Computational and Applied Mathematics*, 146:127–39, 2002.

Method	Computational cost
non-DDM	$O(N_n^p) + O(N_e^p)$
TMSDM on 1 computer	$M_F d^{1-p} O(N_n^p) + d^{1-p} O(N_e^p)$
TMSDM on d computers	$M_F d^{1-p} O(N_n^p) + d^{-p} O(N_e^p)$
TMSDM on d^2 computers	$M_F d^{-p} O(N_n^p) + d^{-p} O(N_e^p)$

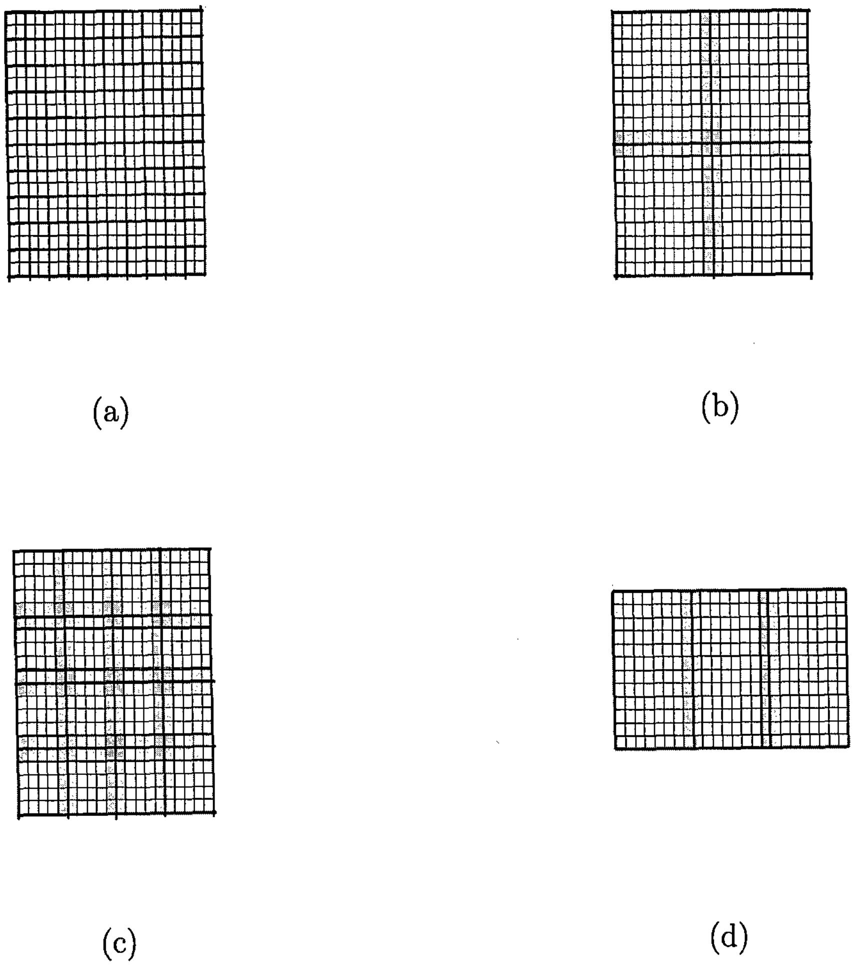
Table 1. The comparison of the computational costs for the non-DDM method, TMSDM for one, d, and d^2 computers.

Method\ Iteration	1	2	3	4	5	6
one-level	.318e-01	.148e-02	.555e-03	.143e-03	.235e-04	.254e-05
two-level	.236e-01	.466e-03	.717e-05	.104e-06	.251e-08	.831e-10

Table 2. The comparison of L^2 error with respect to the LINPACK solution, which is shown in Figure 2(b), for one-level and two-level MODDM, as iteration increases.

Subdomains \ Iteration	1	2	3	4	5	6
2 × 2 subdomains, width 1	.236e-01	.466e-03	.717e-05	.104e-06	.251e-08	.831e-10
2 × 2 subdomains, width 2	.248e-02	.306e-05	.553e-07	.204e-09	.478e-12	.529e-14
2 × 2 subdomains, width 3	.195e-03	.218e-06	.101e-09	.571e-13	.167e-15	.163e-15
4 × 4 subdomains, width 1	.237e-01	.482e-03	.808e-05	.208e-06	.894e-08	.185e-09
4 × 4 subdomains, width 2	.329e-02	.106e-04	.503e-06	.209e-08	.351e-10	.327e-12
4 × 4 subdomains, width 3	.706e-03	.231e-05	.440e-08	.136e-10	.107e-13	.181e-15

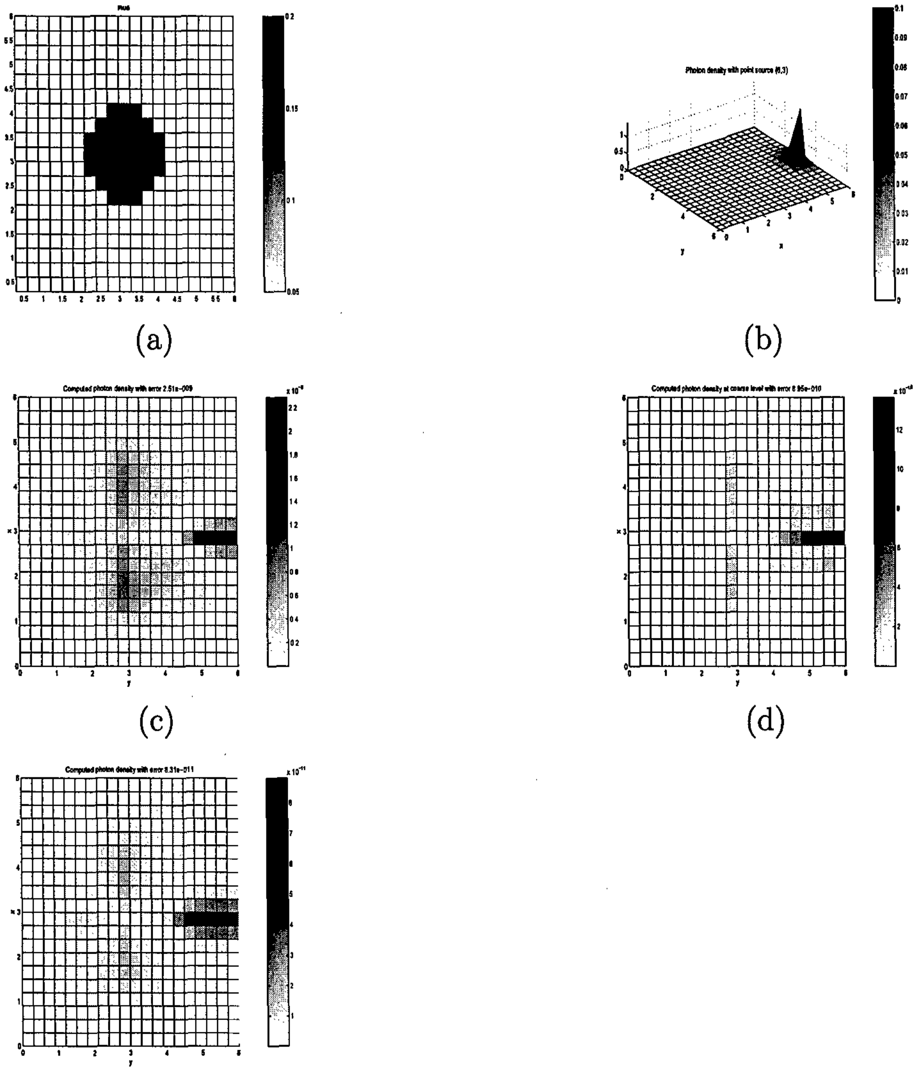
Table 3. The effect of multiple subdomains and overlapping width to the convergence.



(a) (b)

(c) (d)

Figure 1. (a) The lattice of thin lines is 20×20 mesh of $\Omega = [0,6] \times [0,6]$, and thick lines represents the coarse level of Ω, (b) 2×2 domain decomposition of $\Omega = [0,6]\times[0,6]$ with 20×20 mesh, (c) 4×4 domain decomposition of $\Omega = [0,6] \times [0,6]$ with 20×20 mesh, (d) 3×1 domain decomposition of $\Omega = [0,9.6] \times [0,4.8]$ with 24×12 mesh.



(a)

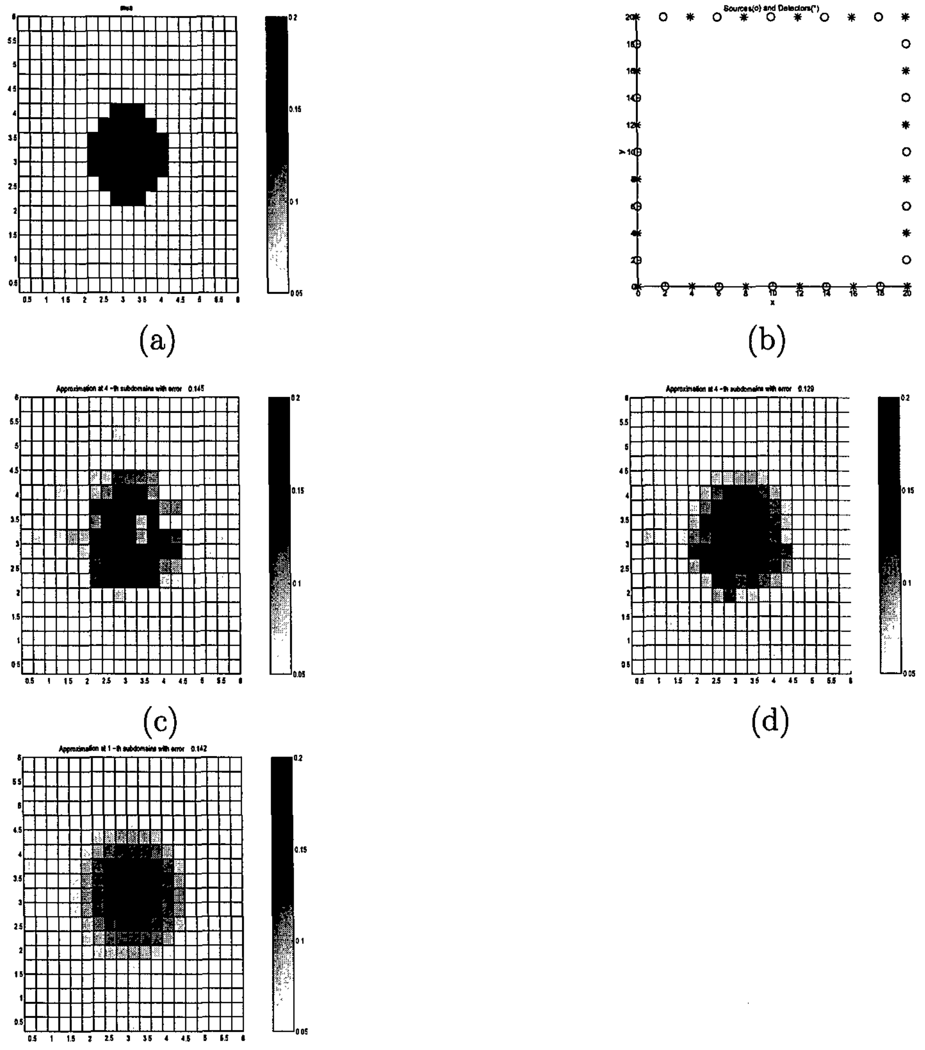
(b)

(c)

(d)

(e)

Figure 2. (a) is the image of μ_a. The white region represent background tissue with value $0.05cm^{-1}$ and the black region represent anomaly which have $0.2cm^{-1}$ absorption coefficient. The photon density with point source at $(6,3)$ solved by LINPACK is depicted in (b). The error between LINPACK solution and the result after 5th iteration, coarse level after 5th iteration, and 6th iteration attained from Algorithm TMODDM is presented in (c), (d), and (e), respectively.

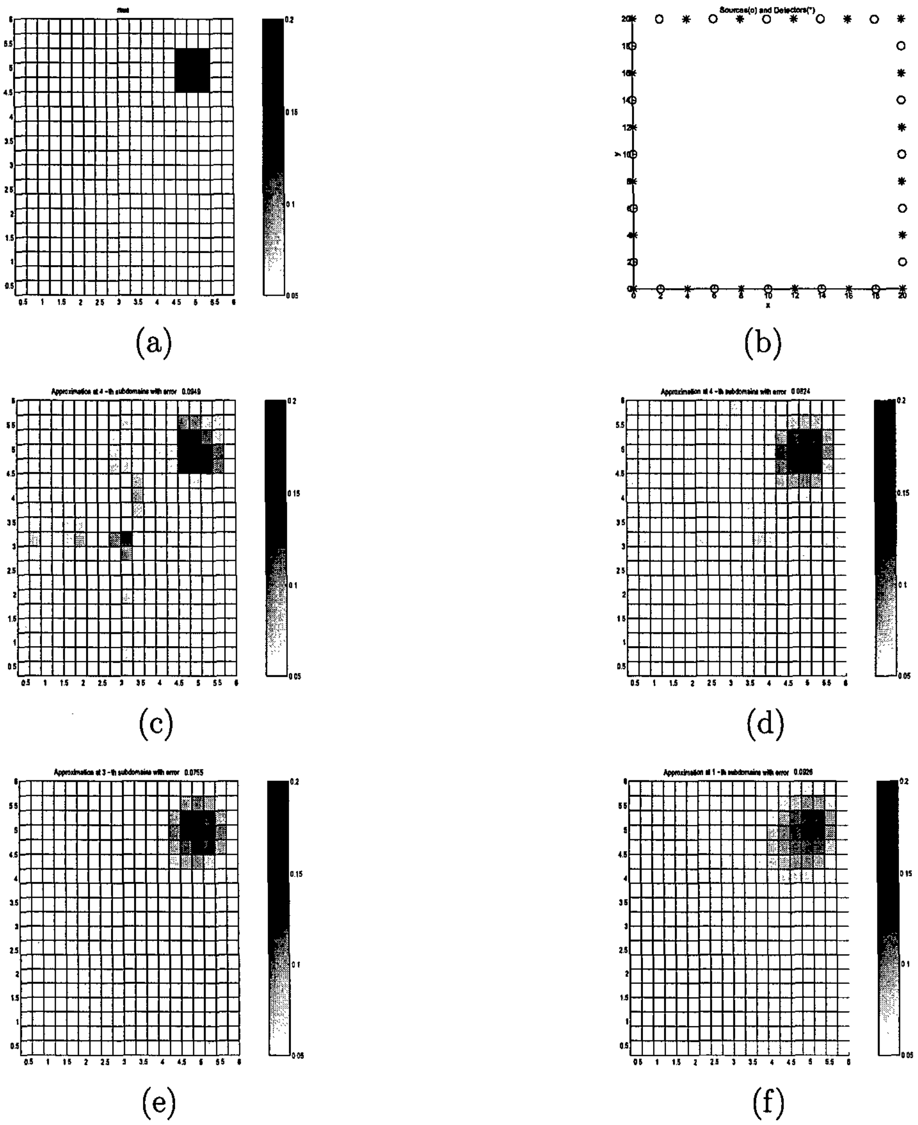


(a) (b)

(c) (d)

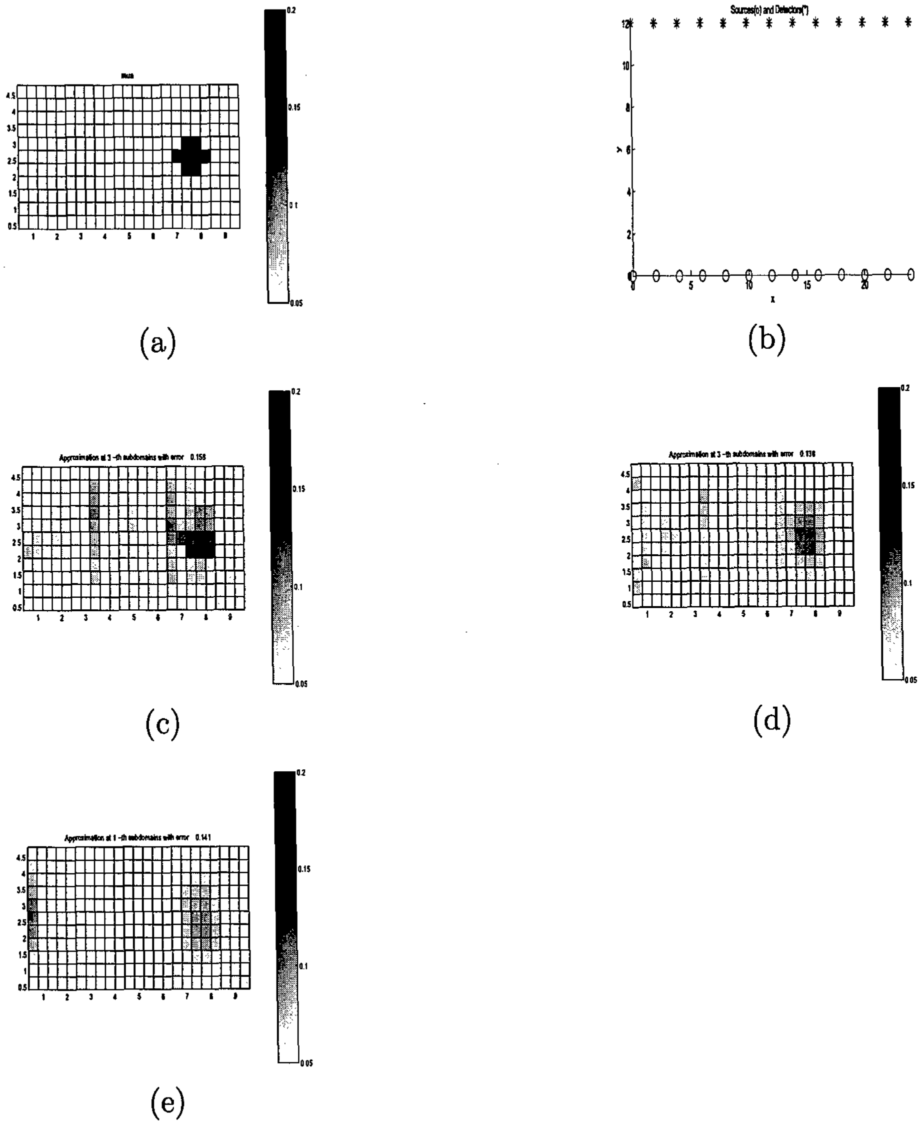
(e)

Figure 3. The reconstruction of absorption coefficient using Algorithm TMSDM on $\Omega = [0,6] \times [0,6]$ with 2×2 subdomains, 20 detectors, 20 sources, and 400 pixels (a)The image of μ_a to be reconstructed. White region represent background tissue with $\mu_a = 0.05cm^{-1}$ and black region represent anomaly with $\mu_a = 0.2cm^{-1}$ (b)The location of 20 sources and 20 detectors (c)The reconstructed image of the absorption coefficient with no regularization using Algorithm TMSDM resulting in L^2 norm error 0.145 (d)The reconstructed image of the absorption coefficient with $\alpha = 10^{-2}$ using Algorithm TMSDM resulting in L^2 norm error 0.129 (e)The reconstructed image of the absorption coefficient with $\alpha = 10^{-2}$ for non-DDM resulting in L^2 norm error 0.142



(a) (b) (c) (d) (e) (f)

Figure 4. The reconstruction of the absorption coefficient on $\Omega = [0,6] \times [0,6]$ with 2×2 subdomains, 20 detectors, 20 sources, and 400 pixels. (a) The image of μ_a to be reconstructed. White region represent background tissue with $\mu_a = 0.05cm^{-1}$ and black region represent anomaly with $\mu_a = 0.2cm^{-1}$. (b) The location of 20 sources and 20 detectors (c) The reconstructed image of the absorption coefficient using Algorithm TMSDM with $\alpha = 0$ resulting in L^2 norm error 0.0949 (d) The reconstructed image of the absorption coefficient using Algorithm TMSDM with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.0824 (e) The reconstructed image of the absorption coefficient using Algorithm TMSDM only on fourth subdomain with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.0755 (f) The reconstructed image of the absorption coefficient for non-DDM with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.0926



(a) (b) (c) (d) (e)

Figure 5. The reconstruction of absorption coefficient on $\Omega = [0, 9.6] \times [0, 4.8]$ with 3×1 subdomains, 13 detectors, 13 sources, and 288 pixels. (a) The image of μ_a to be reconstructed. White region represent background tissue with $\mu_a = 0.05cm^{-1}$ and black region represent anomaly with $\mu_a = 0.2cm^{-1}$. (b) The location of 13 sources and 13 detectors (c) The reconstructed image of the absorption coefficient using Algorithm TMSDM with $\alpha = 0$ resulting in L^2 norm error 0.156. (d) The reconstructed image of the absorption coefficient using Algorithm TMSDM with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.138. (e) The reconstructed image of the absorption coefficient for non-DDM with $\alpha = 10^{-2}$ resulting in L^2 norm error 0.141.

A 2-LEVEL DOMAIN DECOMPOSITION ALGORITHM FOR INVERSE DIFFUSE OPTICAL TOMOGRAPHY

Il-Young Son, Murat Guven, Birsen Yazici

Rensselaer Polytechnic Institute
Electrical, Computer, and Systems
Engineering Troy, NY, USA

Xavier Intes

University of Pennsylvania
Department of Biophysics and Biochemistry
Philadelphia, PA, USA

ABSTRACT

In this paper, we explore domain decomposition algorithms for the inverse DOT problem in order to reduce the computational complexity and accelerate the convergence of the optical image reconstruction. We propose a combination of a two-level multigrid algorithm with a modified multiplicative Schwarz algorithm, where a conjugate gradient is used as an accelerator to solve each sub-problem formulated on each of the partitioned sub-domains. For our experiments, simulated phantom configuration with two rectangular inclusions is used as a testbed to measure the computational efficiency of our algorithms. No a priori information about the configuration is assumed except for the source and detector locations. For the application of our modified Schwarz algorithm alone, we observe an increase in efficiency of 100% as compared to the conjugate gradient solution obtained for the full domain. With the addition of the coarse grid, this efficiency rises to 400%. The coarse grid also serves to improve the overall appearance of the reconstructed image at the boundaries of the inclusions.

1. INTRODUCTION

Inverse diffuse optical tomography (DOT) problem involves estimation of the optical properties of biological tissues, which are pertinent to tissue's physiological and biochemical state. The most prominent applications of DOT are in detecting tumors in the breast, monitoring brain activity, and detecting brain tumors and hemorrhages.

DOT poses a computationally challenging inverse problem. Thereby, realization of real-time diffuse optical tomographic imaging requires computationally viable reconstruction algorithms that provide accurate quantitative results. In this work, we address the computational complexity of the inverse problem by proposing a two-level domain decomposition procedure. Domain decomposition methods convert the inverse problem into smaller-size problems that are easier to handle. Herein we apply a modified alternating Schwarz method with conjugate gradient (CG) algorithm as the iterative solver to accelerate the solution. We then extend the uni-level problem formulation to a 2-level problem to include a coarse grid correction in an attempt to improve quantitative accuracy and further accelerate the convergence.

The Schwarz algorithm begins by partitioning the domain into two or more overlapping sub-domains. The problem is then divided into subproblems on each of these sub-domains. Multiplicative Schwarz algorithm successively solves the localized problems in each sub-domain, every time using the latest solution available to initialize the next sub-problem. Therefore, at every iteration one sub-problem, initialized with the solution from the last iteration, is solved, whose solution is then used to initialize the next sub-domain and so on. We follow the multiplicative Schwarz algorithm using a fixed number of CG iteration to approximate the solution for each sub-domain. For full explanation of Schwarz algorithms see [1, 2].

Single level Schwarz algorithms are not well suited for problems exhibiting low frequency errors and suffer from low convergence rate. Multi-grid methods, on the other hand, use a hierarchy of grids at different scales to accelerate the convergence of standard iterative methods. The fundamental idea behind all multi-grid methods is to combine computations done on different grid scales in order to eliminate the error components of the finest grid, where the original problem has been formulated. This is achieved by approximating the *smoothed* fine grid error on a coarser grid where it can be accurately represented. The approximated error on the coarse grid appears to be more oscillatory, hence can be further eliminated by the iterations on the coarse grid. The solution obtained for the error on the coarse grid is then interpolated to the fine grid to correct the current fine grid solution estimate. There is no unique way of formulating an inverse problem in a multigrid framework. The formulation needs to be specific to the problem at hand, that is it must be tuned to address the requirements asserted in the problem. Note that, carefully designed multi-grid solvers have the potential of solving inverse problems with N unknowns within $O(N)$ work load [3], which makes them the most efficient solvers for many kinds of mathematical problems.

Multigrid methods have been applied for the DOT problem in the past, where the inverse problem is formulated on a hierarchy of *regular* rectangular grids [4]. Domain decomposition has also been previously proposed for Bayesian formulation of inverse DOT problem [5]. However, the nature of the decomposition was different in that work, where non-overlapping, and hence non-Schwarz type, "sliding window" decomposition was used. In our previous work, we have proposed Fast Adaptive Composite-grid (FAC) algorithms for Region-of-Interest DOT [6, 7], which implicitly pursued a domain decomposition with the aid of *a priori* information about the medium of interest. Note that FAC can be viewed as a Schwarz-like domain decomposition method in terms of fully overlapping sub-domains, hence achieving fast convergence with low cost by the use of coarse grid with substantially fewer points in the overlap region [8].

In this work, we assume that no a priori information is available about the unknown image. We have used the location of sources and detectors to determine the overlapping sub-regions. We formulate the inverse problem on the fine grid as two smaller size inverse problems on overlapping sub-regions. CG accelerator is used to approximate the solution on one sub-region. This solution is used to initialize the optimization in the other sub-region. The fine grid iterations are followed by a coarse grid correction scheme, where the inverse problem aims to solve the residual equation formulated on the global coarse grid, which is of relatively smaller size. A number of 2-grid cycles are run to further improve the accuracy of the reconstruction.

2. PROBLEM FORMULATION

2.1. Forward Model

The forward model for DOT is based on simplifying assumptions applied to radiative transport equation which results in a form of photon diffusion equation. In frequency domain, the diffusion equation is given by:

$$-D\nabla^2\Phi(r) - \frac{j\omega}{c}\Phi(r) + \mu_a(r)\Phi(r) = S \tag{1}$$

where D is the diffusion coefficient, c is the speed of light and $\mu_a(r)$ on the entire domain $\Omega \subset R^2$ is the spatially varying absorption coefficient. S stands for the point source located at $r = r_s$. In this work, we focus on the reconstruction of absorption coefficients μ_a of the medium, hence we assume that the diffusion coefficient D has a spatially uniform distribution. We have employed the perturbation approach [9] around a spatially invariant optical background with a first order Rytov approximation to solve the inverse problem. The cell-centered discretization on the grid Ω^h, yields a system of linear equations relating the differential absorption coefficients $\delta\mu_a(r_l)$ to the measurements:

$$\begin{bmatrix} y_{11}^{f_1} \\ \vdots \\ y_{1m}^{f_1} \\ \vdots \\ y_{11}^{f_2} \\ \vdots \\ y_{nm}^{f_p} \end{bmatrix} = \begin{bmatrix} W_{111}^{f_1} & \cdots & W_{11N}^{f_1} \\ \vdots & \ddots & \vdots \\ W_{1m1}^{f_1} & \cdots & W_{1mN}^{f_1} \\ \vdots & \ddots & \vdots \\ W_{111}^{f_2} & \cdots & W_{11N}^{f_1} \\ \vdots & \ddots & \vdots \\ W_{nm1}^{f_p} & \cdots & W_{nmN}^{f_p} \end{bmatrix} \times \begin{bmatrix} \delta\mu_a(r_1) \\ \vdots \\ \vdots \\ \delta\mu_a(r_N) \end{bmatrix} \tag{2}$$

where $y_{ij}^{f_k}$ denote the real part of the measurement at i^{th} source and j^{th} detector at frequency f_k, $W_{ijh}^{f_k}$ is the weight for the lth pixel for ij source-detector pair, and $\delta\mu_a(r_l)$ is the differential absorption coefficient for l_{th} pixel. We can denote this model system succinctly as:

$$y = W^h x^h \tag{3}$$

where W^h is the weight matrix and $y \in \mathbb{R}^M$, $x \in \mathbb{R}^N$, $W^h \in \mathbb{R}^{M\times N}$. N is the number of grid points on Ω^h and M is the total number of measurements.

2.2. Inverse Problem

We formulate the discrete inverse problem to yield a minimum least squares solution for the differential absorption coefficients x^h on Ω^h.

$$\hat{x}_{LS}^h = \arg\min_{x^h} \Psi(x^h) = \arg\min_{x^h} \|y - W^h x^h\|^2 \tag{4}$$

where $||\cdot||$ denotes the Euclidean norm.

2.2.1. *Uni-Level Domain Decomposition Algorithm*

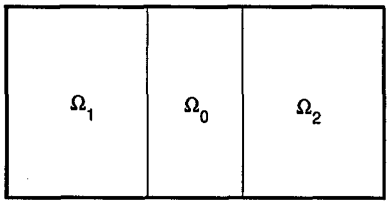


Fig. 1. Decomposition of Ω

In this work, we decompose the domain Ω into two overlapping sub-domains Ω_1 and Ω_2 such that $\Omega = \Omega_1 \cup \Omega_2$ and Ω_0 is the overlapping region, that is $\Omega_0 = \Omega_1 \cap \Omega_2$ (Figure 1). We discard those source-detector pairings that would produce a coupling between the two sub-domains outside the overlapping region. In other words we remove any measurement that is due to a source in $\Omega_1\backslash\Omega_2$ and to a detector in $\Omega_2\backslash\Omega_1$ or vice versa. As a result, the measurement vector $\tilde{y} \in R^{\tilde{M}}$ for $\tilde{M} \leq M$ becomes

$$\tilde{y} = \begin{pmatrix} y_{\Omega_1\backslash\Omega_2} \\ y_{\Omega_0} \\ y_{\Omega_2\backslash\Omega_1} \end{pmatrix} \tag{5}$$

Regrouping the measurements in two vectors y_1 and y_2 yields $y_1 = (y_{\Omega_1\backslash\Omega_2} | \, y_{\Omega_0})^T$ and $y_2 = (y_{\Omega_0} | \, y_{\Omega_2\backslash\Omega_1})^T$. Similarly the differential absorption coefficients on the sub-domains are grouped as x_1^h and x_2^h, such that x_1^h is a finite dimensional approximation of differential absorption coefficients on Ω_1^h and x_2^h on Ω_2^h. As a result we can formulate two sub-domain problems as follows

$$\begin{aligned} \hat{x}_{1,LS}^h &= \arg\min_{x_1^h} \|y_1 - W_1^h x_1^h\|^2 \\ \hat{x}_{2,LS}^h &= \arg\min_{x_2^h} \|y_2 - W_2^h x_2^h\|^2 \end{aligned} \tag{6}$$

Instead of estimating x^h using the formulation given in equation 4, we propose a procedure that follows minimization of the two objective functionals formulated on the two sub-domains, with reduced number of measurements. This results in significant reduction in the size of the inverse problem and consequently the computational complexity of the overall inverse DOT problem.

A conjugate gradient (CG) algorithm is utilized to accelerate the solution on each sub-grid. The solution update obtained after a sweep of CG iterations on one domain is restricted to the overlapping region on the grid Ω_0^h by the discrete operator $\mathcal{I} : \Omega_1^h \rightarrow \Omega_0^h$. The restricted estimates are then used to update the x_2^h estimates on Ω_0^h, which is followed by iterations on Ω_2^h to yield a solution on entire discrete domain Ω_2^h. A υ number of cycles is applied until a desired level of convergence is achieved.

This approach takes advantage of the reduced size of inverse problem formulation by decomposing the domain and the associated measurements. Initialization in one sub-problem solution by the current estimates in the other one and use of CG algorithm are other important factors facilitating the acceleration of the solution.

2.2.2. *Integrating Coarse Grid Correction*

Single level Schwarz algorithms are not well suited for problems exhibiting low frequency errors. CG iterations tend to *smooth* the error in the solution. Even though CG algorithm accelerates the convergence to the solution x^h on Ω^h, it is unable to further eliminate the error with smooth content. This necessitates the use of a coarse grid correction scheme, which enables elimination of smooth error by restricting it onto a coarser grid and correcting the fine grid solution by interpolating the error in the coarse grid.

After the end of CG cycles on the sub-problems, the current estimates x_1^h and x_2^h are concatenated to yield the overall solution estimate x^h on Ω^h. The error between the actual solution x^* and the current solution estimate $\hat{x}^h$ is given by $e^h = x^* - \hat{x}^h$. Assuming that the error e^h on Ω^h is smoothed well enough by CG iterations, we can write a coarse grid approximation for this error, such that

$$e^h = I_{2h}^h e^{2h}$$

where I_{2h}^h is the interpolation operator. We have selected $I_{2h}^h : \mathbb{R}^{N/4} \longrightarrow \mathbb{R}^N$ as $4(I_h^{2h})^T$ where $I_h^{2h} : \mathbb{R}^N \longrightarrow \mathbb{R}^{N/4}$ is the full weighting operator in 2D [10]. As a result, the objective functional given in equation 4 can be re-written as

$$\begin{aligned} \|y - W^h x^h\|^2 &= \|y - W^h(\hat{x}^h + e^h)\|^2 \\ &= \|y - W^h(\hat{x}^h + I_{2h}^h e^{2h})\|^2 \\ &= \|y - W^h\hat{x}^h - W^h I_{2h}^h e^{2h})\|^2 \\ &= \|r - W^{2h} e^{2h}\|^2 \qquad (7) \end{aligned}$$

which is the coarse grid objective functional defined on the coarse grid Ω^{2h} with grid size of $2h$. $r = y - W^h\hat{x}^h$ is called the residual, e^{2h} is the coarse grid error, and $W^{2h} = W^h I_{2h}^h$ is the coarse grid operator. CG iterations on the coarse grid error e^{2h}, which is initially assigned to zero, eliminate the high-frequency components of e^{2h} which appeared to be smooth on Ω^h. This results from the fact that low frequency components appear more oscillatory on coarse grids as compared to fine grids.

The error estimate e^{2h} on Ω^{2h} can then be interpolated to the fine grid to correct the fine grid solution estimate x^h

$$x^h \leftarrow x^h + I_{2h}^h e^{2h} \qquad (8)$$

A predefined number of 2-grid cycles are run to further improve the solution accuracy. A pseudo-code of the overall 2-level domain decomposition algorithm is given below.

Algorithm 1 Two-Grid Domain Decomposition Algorithm

```
$W_1^h, W_2^h, y_1, y_2 \leftarrow partition(W^h, y)$
$W^{2h} \leftarrow restrict(W^h)$ {We only need to generate $\Omega^{2h}$ once}
$x^h \leftarrow initialize(x^h)$
repeat
    $x_1^h \leftarrow CG(W_1, y_1, x_1)$ {Smoothing on $\Omega_1^h$}
    $x_2^h \leftarrow CG(W_2, y_2, x_2)$ {Smoothing on $\Omega_2^h$}
    $r \leftarrow y - W^h x^h$ {Calculate residual.}
    $e^{2h} \leftarrow CG(W^{2h}, r, e^{2h})$ {Solve for coarse grid error}
    $e^h \leftarrow interpolate(e^{2h})$
    $x^h \leftarrow x^h + e^h$
until convergence
```

3. EXPERIMENTAL RESULTS

We performed test cases against a simple simulated phantom configuration consisting of two rectangular inclusions as seen in Fig. 3. The resolution of our test image was 20 pixels by 40 pixels and the number of sources and detectors used was 17 and 33 respectively resulting in a 561 by 800 weight matrix W^h. We compared four test cases: (1) least squares solution using conjugate gradient on Ω^h for the full inverse problem, (2) least squares solution using conjugate gradient on Ω^h with the reduced source and detector configurations, (3) uni-level domain decomposition on the reduced source and detector configuration, and (4) two-level domain decomposition with modified multiplicative smoother on the reduced source and detector configurations. By discarding those measurements that couple source on one sub-domain and the detector on the other, we effectively reduce the dimension of the weight matrix W^h.

All algorithms were implemented in MATLAB. The results were compared using the same conjugate gradient and decomposition codes. Figure 2 shows a plot of square error between the actual and estimated image versus floating point operations (flops) required for reconstruction. The square error was calculated by taking the pixel by pixel difference between the true image (see Figure 3) and the reconstructed image then taking the sum of the squares of those differences. As shown in Figure 2, there is up to an average of 100% increase in efficiency using the uni-level domain decomposition algorithm compared with case (1). With the addition of coarse grid correction, there is approximately 400% increase in efficiency. We observed that the error curve tends to settle faster for single level methods. This is to be expected since the motivation for coarse grid correction is to compensate for the smoother's inability to properly handle globalized (i.e. low frequency) errors.

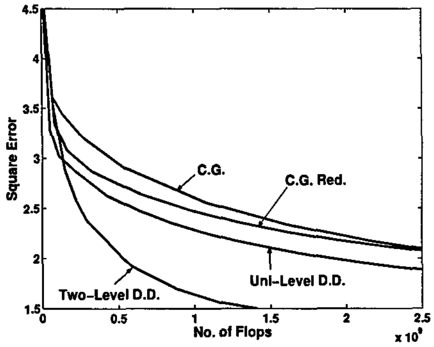


Fig. 2. Square Error vs. No. of Flops

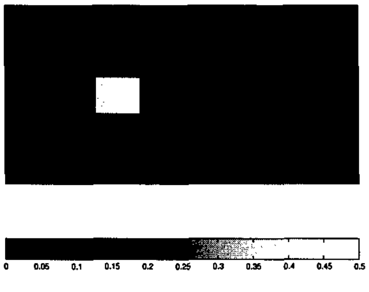


Fig. 3. Simulated Phantom Configuration

Figures 4, 5, 6, and 7 shows the reconstructed images for each method after 1200 iterations. As can be inferred from Figures 5

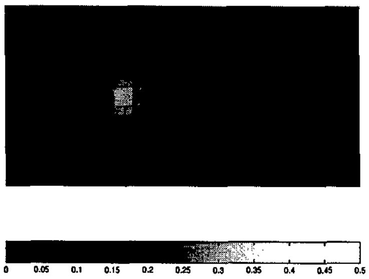


Fig. 4. Direct Method: 1200 Iterations

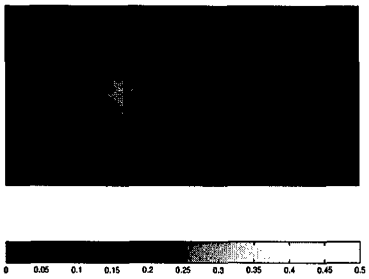


Fig. 5. Direct Method w/ Reduced S-D Pairs: 1200 Iterations

and 6, the uni-level domain decomposition method and the direct CG on the reduced source-detector configurations had nearly equal error after 1200 iterations, but the computational time was cut by half for the domain decomposition case. The effect of coarse grid correction on the domain decomposition algorithm is clear from Figure 7. By incorporating coarse grid correction, the algorithm is able to handle the smooth errors around the boundaries of the inclusions. In effect, we end up with a much sharper picture and better results around these boundaries.

4. CONCLUSION

In this paper, we investigated the effectiveness, in terms of computational efficiency, of applying Schwarz type domain decomposition method in the solution of linearized DOT inverse problem. We modified the classic Schwarz algorithm where CG was used to accelerate the convergence to the least squares solution for each sub-problem in the place of the usual Gauss-Siedel iterations. In our two grid domain decomposition algorithm, CG was applied to each sub-domain for several iterations (to get a good smoothing effect). The residual from the entire domain was then restricted to the coarse grid to find the coarse grid error which was then interpolated back to the fine grid for error correction. The addition of the domain decomposition reduced the computational burden while addition of the coarse grid allowed for correction of global (low frequency) errors. In terms of computational efficiency as compared with CG solution for the full problem, uni-level domain decomposition method showed 100% increase while two-level method saw nearly 400%. As a result of these preliminary findings, we feel that domain decomposition coupled with multi-grid methods is a viable option for decreasing the computation complexity of the DOT inverse problem even when there is no parallelization of the problem. As such, the algorithm proposed is well-suited for real-time DOT image reconstruction.

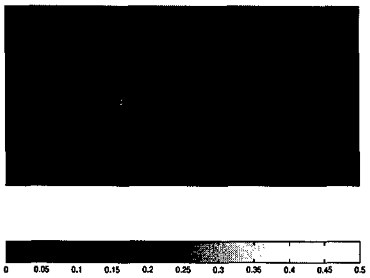


Fig. 6. One Grid D.D.: 1200 Iterations

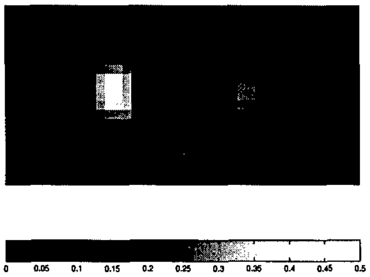


Fig. 7. Two Grid D.D.: 1200 Iterations

5. REFERENCES

[1] Barry Smith, Petter Bjørstad, and William Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, 1996.

[2] Alfio Quarteroni and Alberto Valli. *Domain Decomposition Methods for Partial Differential Equations.* Oxford Science Publications, 1999.

[3] U. Trottenberg, C. Oosterlee, A. Schüller, "Multigrid," Academic Press, New York, 2001.

[4] J. C. Ye, C. A. Bouman, K. J. Webb, and R. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography," IEEE Transactions in Image Processing. Vol. 10, No. 6, pp. 909-922, June 2001.

[5] M. J. Eppstein, D. E. Dougherty, D. J. Hawrysz, E. M. Sevick-Muraca, "Three-Dimensional Bayesian Optical Image Reconstruction with Domain Decomposition," IEEE Transactions on Medical Imaging, Vol. 20, No. 3, March 2001.

[6] M. Guven, B. Yazici, X. Intes, B. Chance,"An Adaptive Multigrid Algorithm for Region of Interest Diffuse Optical Tomography," Proceedings of IEEE Int. Conference in Image Processing, 2003.

[7] M. Guven, B. Yazici, X. Intes, B. Chance,"An Adaptive V-grid Algorithm for Diffuse Optical Tomography," Proceedings of IEEE Bioengineering Conference, pp: 95 -96, 2003.

[8] S. F. Cormick, "Multilevel Adaptive Methods for Partial Differential Equations," SIAM, Philadelphia, 1989.

[9] S. R. Arridge, "Photon Measurement Density Functions. Part 1: Analytical Forms," *Appl. Opt.* **34**, pp. 7395-7409, 1995.

[10] William L. Briggs, Van Emden Henson, and Steve F. McCormick. *A Multigrid Tutorial: Second Edition.* SIAM, 2000.

INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY
Phys. Med. Biol. **50** (2005) 2837–2858 doi:10.1088/0031-9155/50/12/008

Diffuse optical tomography with *a priori* anatomical information

Murat Guven[1], Birsen Yazici[1], Xavier Intes[2,3] and Britton Chance[2]

[1] Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY, USA
[2] Department of Biophysics and Biochemistry, University of Pennsylvania, Philadelphia, PA, USA

E-mail: yazici@ecse.rpi.edu

Received 4 February 2005, in final form 6 February 2005
Published 1 June 2005
Online at stacks.iop.org/PMB/50/2837

Abstract
Diffuse optical tomography (DOT) poses a typical ill-posed inverse problem with a limited number of measurements and inherently low spatial resolution. In this paper, we propose a hierarchical Bayesian approach to improve spatial resolution and quantitative accuracy by using *a priori* information provided by a secondary high resolution anatomical imaging modality, such as magnetic resonance (MR) or x-ray. In such a dual imaging approach, while the correlation between optical and anatomical images may be high, it is not perfect. For example, a tumour may be present in the optical image, but may not be discernable in the anatomical image. The proposed hierarchical Bayesian approach allows incorporation of partial *a priori* knowledge about the noise and unknown optical image models, thereby capturing the function-anatomy correlation effectively. We present a computationally efficient iterative algorithm to simultaneously estimate the optical image and the unknown *a priori* model parameters. Extensive numerical simulations demonstrate that the proposed method avoids undesirable bias towards anatomical prior information and leads to significantly improved spatial resolution and quantitative accuracy.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Diffuse optical tomography (DOT) is a non-invasive imaging modality that makes use of the light in the near-infrared (NIR) spectrum (Yodh and Chance 1995, Hebden *et al* 1997, Arridge and Hebden 1997, Intes and Chance 2005). The inverse problem in DOT involves

---

[3] Present address: ART Advanced Research Technologies, Quebec, Canada.

0031-9155/05/122837+22$30.00 © 2005 IOP Publishing Ltd Printed in the UK

reconstruction of spatially varying absorption and scattering properties (O'Leary 1996, Boas *et al* 2001, Arridge 1999) as well as fluorophore lifetime and yield (Chang *et al* 1997, Hawrysz and Sevick-Muraca 2000, Eppstein *et al* 2002) in tissues from boundary measurements. These fundamental quantities can be utilized to obtain tissue oxy- and de-oxyhaemoglobin concentrations, blood oxygen saturation, water, fat and contrast agent uptake in tissue (Kincade 2004). The unique physiological and biochemical information offered by DOT is very valuable for practical applications such as breast cancer diagnosis (Cerussi *et al* 2001, Srinivasan *et al* 2003, Intes *et al* 2003), cognitive activity monitoring (Strangman *et al* 2002, Villringer and Chance 1997, Chance *et al* 2003), brain tumour and haemorrhage detection (Hebden *et al* 2004), functional muscle imaging (Quaresima *et al* 2003) with a growing list of applications in fluorescence tomographic imaging (Frangioni 2003, Weissleder and Ntziachristos 2003).

DOT poses a typical ill-posed inverse problem with a large number of unknowns and a relatively limited number of measurements. This necessitates the incorporation of *a priori* information into the inverse problem formulation in order to obtain viable solutions. Furthermore, propagation of NIR light is not restricted to a plane owing to the diffuse nature of photons in turbid media, which results in poor spatial resolution. To tackle the ill-posed nature of the inverse problem and to address the low spatial resolution in DOT, a number of approaches have been developed. Bayesian approach has been suggested to incorporate *a priori* information to the inverse problem formulation (Oh *et al* 2002, Milstein *et al* 2002, Eppstein *et al* 2002, Guven *et al* 2002, Ye *et al* 2001). Introducing penalty functions (Hielscher and Bartel 2001) and uniform (Paulsen and Jiang 1996, Arridge 1993, Jiang *et al* 1996, Yao *et al* 1997) or spatially varying regularization terms (Pogue *et al* 1999) within the regularization framework are alternative ways to incorporate *a priori* information into the image reconstruction process. In all these studies, no other imaging modality has been utilized to infer information specific to the medium of interest, which could be used to tune the prior information.

1.1. Related literature

Recently several research groups reported development of hybrid imaging systems combining optical methods with high resolution anatomical imaging techniques. These include a concurrent x-ray tomosynthesis-DOT system at Massachusetts General Hospital (Li *et al* 2003), MRI-DOT/DOS (Diffuse Optical Spectroscopy) systems at University of Pennsylvania (Intes *et al* 2002), University of California at Irvine (Gulsen *et al* 2003) and Dartmouth College (Brooksby *et al* 2003) and ultrasound-DOT system at University of Connecticut (Zhu *et al* 2003a). These multi-modality developments are all motivated by the fact that DOT offers unique functional information (such as tissue oxy- and deoxy-haemoglobin concentrations) while high resolution anatomical imaging modalities provide complementary information for disease diagnosis and understanding with superior localization and spatial resolution. Another incentive comes from the assumption that the contrast elements provided by high resolution imaging modalities correlate well with the optical properties. A number of studies lend support to this assumption. Ntziachristos *et al* (2000) have reported that there exists a good spatial correlation between gadolinium (Gd)-enhanced MR and Indocyanine Green (ICG)-enhanced DOT images. Cuccia *et al* (2003) have also shown that Gd-enhanced and Methyl Blue (MB)-NIR results correlate well with each other in terms of perfusion dynamics. Merritt *et al* (2003a, 2003b) presented similar observations to demonstrate correlation between MR and DOS in water and lipid concentration retrieval. Furthermore a number of studies have shown that incorporation of high resolution anatomical images as *a priori* information leads to improved diffuse optical image reconstruction (Dehghani *et al* 2002, Ntziachristos *et al*

2002). Pogue and Paulsen (1998) used MR images to generate a finite element mesh to reconstruct a simulated rat cranium, where the available information, about the structure and optical properties, is used for the initial guess in the inversion algorithm similar to the approach followed by Xu *et al* (2002). Schweiger and Arridge (1999) suggested using the structural information to reconstruct images of a segmented brain model at a low resolution level in order to obtain a good initial guess for the high-resolution solution of the same problem. Use of MR scans has been employed for optical breast imaging (Chang *et al* 1997, Barbour *et al* 1995) where the 'reference medium' was obtained from accurate optical properties of the tissue, with the anatomy derived from MR images. Li *et al* (2003) reported optical breast imaging results guided by x-ray mammography, where x-ray contrast was assumed to be proportional to DOT contrast. X-ray images are used as the spatial constraint to decompose the optical medium into two major sub-domains, representing the region of interest as referred to the tumour region and the background, respectively. A hybrid L-curve method is followed for the estimation of regularization parameters for each of the regularization terms corresponding to the sub-domains, which challenges the inverse problem computationally. Brooksby *et al* (2003) extended the idea proposed by Schweiger and Arridge (1999) to incorporate the initial low-resolution optical image as derived from MR image by using structural information and spatially varying regularization. The reported results are encouraging; however, accurate quantification of the tumour region is possible only when the true optical heterogeneity of tissue distribution is included. Therefore in this approach, the overall performance relies upon the quality of the initial guess.

1.2. Proposed method

In all the studies referenced above, the performance of the DOT image reconstruction relies on the assumption that the correlation between the anatomical and optical images is high. However, there may be regions in the optical image that do not have any anatomical counterparts. For example, a tumour may be apparent in the optical image, but may not have a corresponding signature in the anatomical image. Furthermore, average optical coefficients extracted from anatomical images may be significantly different from the true optical coefficients of tissue. As a result, the assumption of strong optical-anatomy correlation may cause undesirable, erroneous bias in optical image reconstruction. Therefore, more flexible prior models are needed to properly represent optical-anatomy correlation. For example, when the average optical properties extracted from anatomical images are not reliable, prior image model should provide weaker constraints in image reconstruction. The hierarchical Bayesian framework affords such a flexibility in designing prior image and noise models. In the hierarchical Bayesian framework, one can formulate the inverse problem in multiple stages where each stage includes information about the unknown parameters of the preceding stage. The first stage of the hierarchy includes the data likelihood and the first stage of the image prior, which comprise statistical models for the noise and optical image, respectively. These models include parameters associated with noise and image statistics, which are not known precisely in practice. These unknown parameters are referred to as hyperparameters, which can be regarded as random variables. The succeeding stage of the hierarchical formulation incorporates *a priori* information about the hyperparameters in the form of prior distributions—so called hyperpriors—defined on the hyperparameters. The incorporation of the second stage concludes the design of the two-level hierarchical noise and image models.

In this paper, we consider a two-level hierarchical Bayesian formulation to incorporate *a priori* anatomical and tissue classification information into the DOT image reconstruction.

We start with the segmentation of the high resolution image and classify the segmented image into sub-images representing major tissue types. Based on the tissue label information extracted from the anatomical image, we design the first stage of the prior distribution on the unknown optical image as a function of unknown hyperparameters, namely the image mean and standard deviation. Next, we design the data likelihood corresponding to the parametric noise model with an associated unknown hyperparameter, which is related to the noise variance. The uncertainty in the models owing to the unknown hyperparameters is addressed by defining hyperpriors on the hyperparameters, which constitutes the second stage of the hierarchical formulation. The hyperprior on the noise-variance related hyperparameter is assumed to be uniform so as to not constrain its value. The hyperpriors on the hyperparameters of the image model are formulated with the aid of coregistered tissue classification. Consequently, the second stage of the image prior integrates the subjective information into the formulation, defining the extent of the correlation between the anatomical and optical images. As a result, *a priori* information is used to constrain the hyperparameters, thereby imposing weaker constraints on the optical image. We refer to sections 2 and 3 for a detailed discussion of the hierarchical noise and image models.

Having designed the hierarchical noise and image models, we formulate the joint distribution of the measurements, the image and the unknown hyperparameters associated with the noise and image models. In order to estimate the hyperparameters, we adapt the linear conjugate gradient (CG) algorithm to include a hyperparameter estimation step followed by an image update. In this context, we apply an iterative empirical Bayesian approach to estimate the hyperparameters, which in turn gives the maximum *a posteriori* (MAP) estimates of the hyperparameters at each CG iteration prior to the image update. Hence, the noise and image models are accommodated at each update of the hyperparameters along with the solution process.

We perform simulation experiments to evaluate the performance of the proposed hierarchical Bayesian formulation and hyperparameter estimation scheme. Our results indicate that hierarchical Bayesian approach captures the function-anatomy correlation properly and provides improved DOT image reconstruction without introducing undesirable bias towards *a priori* anatomical information. Our simulation experiments show that the proposed method provides accurate reconstruction of tumours even when tumour contrast is absent in the anatomical image.

1.3. Organization of the paper

The paper is organized as follows. Section 2 defines the forward model. Section 3 provides background on the hierarchical Bayesian formulation of the inverse problem and describes the components of the hierarchical model. Section 4 presents the iterative algorithm for the simultaneous estimation of the optical image and the unknown hyperparameters. Section 5 includes numerical experiments to validate the properties of the proposed approach. Section 6 summarizes our results and conclusion. The appendix includes the derivation of the estimation of the hyperparameters.

2. Forward model

In the NIR region of the electromagnetic spectrum, light propagation in biological tissue can be modelled by the diffusion approximation to the radiative transfer equation. The diffusion equation in the frequency domain is given by

$$\nabla \cdot D(r) \nabla \phi(r) - \mu_a(r)\phi(r) - \frac{\mathrm{i}\omega}{c}\phi(r) = -A\delta(r - r_s), \qquad (1)$$

where $\phi(\boldsymbol{r})$ represents the spatially varying optical field in the medium $\Omega \subset \mathbb{R}^3$, due to the point source $A\delta(\boldsymbol{r}-\boldsymbol{r}_s)$ located at $\boldsymbol{r}=\boldsymbol{r}_s$. ω denotes the modulation frequency of the source, c is the speed of light and $\mathrm{i}=\sqrt{-1}$. $D(\boldsymbol{r})$ is the spatially varying diffusion coefficient and $\mu_a(\boldsymbol{r})$ stands for the spatially varying absorption coefficient of the medium.

In this work, we focus on the reconstruction of absorption coefficients of the medium. Therefore, we assume that the diffusion coefficient of the medium is known and spatially invariant. As a result, the following diffusion equation given in frequency domain suffices to define the forward model:

$$D\,\nabla^2\phi(\boldsymbol{r}) - \mu_a(\boldsymbol{r})\phi(\boldsymbol{r}) - \frac{\mathrm{i}\omega}{c}\phi(\boldsymbol{r}) = -A\delta(\boldsymbol{r}-\boldsymbol{r}_s). \tag{2}$$

We have employed the perturbation approach (O'Leary *et al* 1995, Arridge 1995, Kak and Slaney 1988) with a first-order Rytov approximation to solve the forward problem in the frequency domain to yield a system of linear equations after the discretization of the medium Ω into N uniform voxels (Guven *et al* 2003a):

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x} + \boldsymbol{\zeta}, \tag{3}$$

where $\boldsymbol{y}$ is the measurement vector, $\boldsymbol{W}$ is the Jacobian based on the Rytov approximation computed around a specified homogeneous background $\mu_a(\boldsymbol{r}) = \mu_{a0}$, $\boldsymbol{x} \in \mathbb{R}^N$ denotes the vector of differential absorption coefficients $\delta\mu_a$ of the medium with respect to the homogeneous background and $\boldsymbol{\zeta}$ is the additive noise in the measurement system. Note that recently a number of researches (Ntziachristos *et al* 2002, Li *et al* 2003, Zhu *et al* 2003b, Intes *et al* 2003) have reported improved DOT reconstructions for clinical images based on linearized forward model using high resolution anatomical priors.

3. Hierarchical Bayesian formulation of the inverse problem

We approach the DOT inverse problem from a Bayesian perspective. In particular, we propose a hierarchical Bayesian approach to effectively capture the function-anatomy correlation.

We formulate the posterior distribution of the unknown image and compute its maximum *a posteriori* (MAP) estimate $\hat{\boldsymbol{x}}_{\mathrm{MAP}}$; that is

$$\hat{\boldsymbol{x}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{x}}\{\log p(\boldsymbol{x}|\boldsymbol{y})\},$$

where $\log p(\boldsymbol{x}|\boldsymbol{y})$ is the posterior distribution of the unknown image $\boldsymbol{x}$ given the measurements $\boldsymbol{y}$. Equivalently,

$$\hat{\boldsymbol{x}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{x}}\{\log p(\boldsymbol{y}|\boldsymbol{x}) + \log p(\boldsymbol{x})\},$$

where $p(\boldsymbol{y}|\boldsymbol{x})$ is the data likelihood and $p(\boldsymbol{x})$ is the prior on the unknown image $\boldsymbol{x}$.

In our problem, in addition to the optical boundary measurements $\boldsymbol{y}$, we also have the anatomical tissue label information $\boldsymbol{C}$, derived from the *a priori* anatomical image. Therefore, the MAP estimate can be modified as

$$\begin{aligned}\hat{\boldsymbol{x}}_{\mathrm{MAP}} &= \arg\max_{\boldsymbol{x}}\{\log p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{C})\}\\ &= \arg\max_{\boldsymbol{x}}\{\log p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{C}) + \log p(\boldsymbol{x}|\boldsymbol{C})\},\end{aligned} \tag{4}$$

where $p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{C})$ is the data likelihood and $p(\boldsymbol{x}|\boldsymbol{C})$ is the conditional probability density function of $\boldsymbol{x}$ given the tissue label information $\boldsymbol{C}$.

Given the forward model (3), the data likelihood is governed mainly by the noise statistics. Therefore equation (4) reduces to

$$\hat{\boldsymbol{x}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{x}}\{\log p(\boldsymbol{y}|\boldsymbol{x}) + \log p(\boldsymbol{x}|\boldsymbol{C})\}. \tag{5}$$

In our formulation, the noise statistics and the prior distribution are governed by unknown model parameters λ and $\mathcal{Q}$, respectively. Here, λ is the scalar parameter associated with the noise variance and $\mathcal{Q}$ is the vector of mean and variance of sub-images that correspond to different tissue labels. We will refer to these parameters as hyperparameters. In order to estimate the unknown image and the associated hyperparameters, we consider the joint conditional distribution $p(\boldsymbol{y}, \boldsymbol{x}, \lambda, \mathcal{Q}|\boldsymbol{C})$:

$$p(\boldsymbol{y}, \boldsymbol{x}, \lambda, \mathcal{Q}|\boldsymbol{C}) = p(\boldsymbol{y}, \lambda|\boldsymbol{x})p(\boldsymbol{x}, \mathcal{Q}|\boldsymbol{C}), \tag{6}$$

where $p(\boldsymbol{x}, \mathcal{Q}|\boldsymbol{C})$ is the conditional hierarchical prior on the image $\boldsymbol{x}$ and the hyperparameters $\mathcal{Q}$. Equation (6) can be alternatively expressed as

$$\log p(\boldsymbol{y}, \boldsymbol{x}, \lambda, \mathcal{Q}|\boldsymbol{C}) = \log p(\boldsymbol{y}|\boldsymbol{x}, \lambda) + \log p(\lambda) + \log p(\boldsymbol{x}|\mathcal{Q}, \boldsymbol{C}) + \log p(\mathcal{Q}|\boldsymbol{C}), \tag{7}$$

where $p(\lambda)$ is the prior distribution on λ. We shall refer to $p(\lambda)$ and $p(\mathcal{Q}|\boldsymbol{C})$ as hyperpriors (Berger 1988). In this representation, $p(\boldsymbol{x}|\mathcal{Q}, \boldsymbol{C})$ is the first-stage prior and $p(\mathcal{Q}|\boldsymbol{C})$ stands for the second-stage prior. In the following sections, we will discuss how the data likelihood and the hierarchical prior are modelled.

3.1. The data likelihood model

The measurement vector $\boldsymbol{y}$ is formed as

$$\boldsymbol{y} = \begin{bmatrix} y_{11}^{f_1} & y_{12}^{f_1} & \cdots & y_{1D}^{f_1} & y_{21}^{f_1} & \cdots & y_{SD}^{f_1} & y_{11}^{f_2} & \cdots & y_{SD}^{f_F} \end{bmatrix}^T, \tag{8}$$

where S is the number of sources, D is the number of detectors and F is the number of frequencies associated with each source. The total number of measurements is then equal to $P = S \times D \times F$. For computational efficiency, we limit the data set to the real part of the measurements, thus $\boldsymbol{y} \in \mathbb{R}^P$.

Photon detection can be modelled using shot noise statistics, which originates from Poisson statistics. With a sufficiently large number of detected photons, the Poisson statistics can be approximated by a Gaussian distribution, with a variance proportional to the magnitude of the measurements (Ye *et al* 2001, Oh *et al* 2002). Consequently, we model the data likelihood in equation (6) as

$$p(\boldsymbol{y}, \lambda|\boldsymbol{x}) = \frac{1}{K|\boldsymbol{\Lambda}_\zeta(\lambda)|^{1/2}} \exp\left[-\frac{1}{2}\|\boldsymbol{y} - \boldsymbol{W}\boldsymbol{x}\|^2_{\boldsymbol{\Lambda}_\zeta^{-1}(\lambda)}\right], \tag{9}$$

where we assume a non-informative prior for λ, which is a *uniform density on* $\mathbb{R}^1$ (Berger 1988). In the above distribution, $\boldsymbol{\Lambda}_\zeta(\lambda)$ is the covariance matrix of size $P \times P$, K is the normalization constant and $\|\boldsymbol{z}\|^2_{\boldsymbol{\Lambda}} = \boldsymbol{z}^T\boldsymbol{\Lambda}\boldsymbol{z}$. Under the assumption of statistical independence, $\boldsymbol{\Lambda}_\zeta(\lambda)$ becomes a diagonal matrix of the form:

$$\boldsymbol{\Lambda}_\zeta(\lambda) = \lambda\boldsymbol{\Lambda}_y = \begin{bmatrix} \lambda\sigma^2_{\zeta_1} & 0 & 0 & \cdots & 0 \\ 0 & \lambda\sigma^2_{\zeta_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \lambda\sigma^2_{\zeta_P} \end{bmatrix}, \tag{10}$$

where $\sigma^2_{\zeta_p}$ is equal to the absolute value of the pth measurement and the unknown parameter λ controls the scale of the noise covariance matrix. Therefore, we shall refer to λ as 'noise scale' for the rest of the paper.

3.2. Hierarchical formulation of the prior distribution

We downsample the high resolution anatomical image to match its resolution with that of the optical image. Next, we utilize the downsampled anatomical image to decompose the optical domain into M non-overlapping sub-images, each of which is assumed to represent a major tissue type. For instance, for breast, these tissues types can be parenchyma, adipose and tumour (Ntziachristos *et al* 1999, 2000). We assume a Gaussian distribution for each sub-image, with unknown mean and standard deviation. Thus, the first stage in the hierarchical prior distribution for the ith sub-image is given by

$$p(\boldsymbol{x}_i|\mu_i,\sigma_i) = \frac{1}{\left(2\pi\sigma_i^2\right)^{N_i/2}} \exp\left[-\frac{1}{2\sigma_i^2}\|\boldsymbol{x}_i - \boldsymbol{\mu}_i\|^2\right], \qquad i = 1, 2, \ldots, M \tag{11}$$

with the implicit assumption that the voxels in each sub-image are statistically independent. $\boldsymbol{\mu}_i = (\mu_i \cdots \mu_i)^T$ is the uniform mean value vector of size $N_i \times 1$, where N_i stands for the number of voxels in the ith sub-image. The covariance matrix associated with the ith sub-image $\boldsymbol{\Lambda}_x(\sigma_i) = \sigma_i^2 \boldsymbol{I}_{N_i\times N_i}$ where σ_i is the standard deviation of each voxel in the ith sub-image and $\boldsymbol{I}_{N_i\times N_i}$ is the $N_i \times N_i$ identity matrix. Assuming that the sub-images are statistically independent, the first-stage prior of the image given the tissue label information $\boldsymbol{C}$ is

$$\begin{aligned} p(\boldsymbol{x}|\mathcal{Q},\boldsymbol{C}) &= p(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{C}) \\ &= \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Lambda}_x(\boldsymbol{\sigma})|^{1/2}} \exp\left[-\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Lambda}_x^{-1}(\boldsymbol{\sigma})}\right], \end{aligned} \tag{12}$$

where $\mathcal{Q} = [\boldsymbol{\mu},\boldsymbol{\sigma}]$. $\boldsymbol{\mu}$ is the vector of mean values assigned to the sub-images and $\boldsymbol{\sigma}$ is the vector of standard deviations associated with the sub-images, that is

$$\boldsymbol{\sigma} = \Big[\underbrace{\sigma_1 \cdots \sigma_1}_{N_1} \ \ \underbrace{\sigma_2 \cdots \sigma_2}_{N_2} \ \ \cdots \ \ \underbrace{\sigma_M \cdots \sigma_M}_{N_M}\Big]^T, \tag{13}$$

$$\boldsymbol{\mu} = \Big[\underbrace{\mu_1 \cdots \mu_1}_{N_1} \ \ \underbrace{\mu_2 \cdots \mu_2}_{N_2} \ \ \cdots \ \ \underbrace{\mu_M \cdots \mu_M}_{N_M}\Big]^T, \tag{14}$$

and $\boldsymbol{\Lambda}_x(\boldsymbol{\sigma})$ is the covariance matrix of the image $\boldsymbol{x}$

$$\boldsymbol{\Lambda}_x(\boldsymbol{\sigma}) = \begin{bmatrix} \sigma_1^2 \boldsymbol{I}_{N_1\times N_1} & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 \boldsymbol{I}_{N_2\times N_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sigma_M^2 \boldsymbol{I}_{N_M\times N_M} \end{bmatrix}. \tag{15}$$

The second-stage prior involves incorporation of the *a priori* information into the hierarchical prior distribution; in the form of hyperpriors defined on the unknown hyperparameters of the first stage, that is the mean and standard deviation of different tissue types in the optical image. Note the mean values of different tissue types are specific to the unknown optical image and are different from average optical values of tissues available in the literature. Nevertheless, the information available in the literature can be used to design hyperpriors on the unknown mean and standard deviation, which allows effective modelling of the uncertainty in the prior information.

We assume a Gaussian distribution for the mean value $\boldsymbol{\mu}$ of the image:

$$p(\boldsymbol{\mu}|\boldsymbol{C}) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Lambda}_\mu(\boldsymbol{\vartheta})|^{1/2}} \exp\left[-\frac{1}{2}\|\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}}\|^2_{\boldsymbol{\Lambda}_\mu^{-1}(\boldsymbol{\vartheta})}\right], \tag{16}$$

where $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu}|_{\mu_i=\tilde{\mu}_i}$ and $\tilde{\mu}_i$ is the average differential absorption of the ith tissue type. $\boldsymbol{\Lambda}_\mu(\vartheta) = \boldsymbol{\Lambda}_x(\boldsymbol{\sigma})|_{\sigma_i=\vartheta_i}$ is the covariance matrix where ϑ_i stands for the standard deviation of the mean value of each voxel in the ith tissue type, for $i = 1, 2, \ldots, M$.

Similarly, we assume a Gaussian distribution for standard deviation $\boldsymbol{\sigma}$ of the image:

$$p(\boldsymbol{\sigma}|C) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Lambda}_\sigma(\gamma)|^{1/2}} \exp\left[-\frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\mu}_\sigma\|^2_{\boldsymbol{\Lambda}_\sigma^{-1}(\gamma)}\right], \tag{17}$$

where $\boldsymbol{\mu}_\sigma = \boldsymbol{\mu}|_{\mu_i=\mu_{\sigma_i}}$ and $\boldsymbol{\Lambda}_\sigma(\gamma) = \boldsymbol{\Lambda}_x(\boldsymbol{\sigma})|_{\sigma_i=\gamma_i}$ for $i = 1, 2, \ldots, M$. Thus, the second-stage prior $p(\mathcal{Q}|C)$ becomes

$$\begin{aligned} p(\mathcal{Q}|C) &= p(\boldsymbol{\mu}, \boldsymbol{\sigma}|C) = p(\boldsymbol{\mu}|C)p(\boldsymbol{\sigma}|C) \\ &= \frac{1}{(2\pi)^N|\boldsymbol{\Lambda}_\mu(\vartheta)|^{1/2}|\boldsymbol{\Lambda}_\sigma(\gamma)|^{1/2}} \exp\left[-\frac{1}{2}\left(\|\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}}\|^2_{\boldsymbol{\Lambda}_\mu^{-1}(\vartheta)} + \|\boldsymbol{\sigma} - \boldsymbol{\mu}_\sigma\|^2_{\boldsymbol{\Lambda}_\sigma^{-1}(\gamma)}\right)\right]. \end{aligned} \tag{18}$$

Having designed the first- and second-stage priors, the hierarchical prior distribution in equation (6) becomes

$$\begin{aligned} p(\boldsymbol{x}, \mathcal{Q}|C) &= p(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\sigma}|C) = p(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\sigma}, C)p(\boldsymbol{\mu}|C)p(\boldsymbol{\sigma}|C) \\ &= \frac{1}{(2\pi)^{3N/2}|\boldsymbol{\Lambda}_x(\boldsymbol{\sigma})|^{1/2}|\boldsymbol{\Lambda}_\mu(\vartheta)|^{1/2}|\boldsymbol{\Lambda}_\sigma(\gamma)|^{1/2}} \\ &\quad \times \exp\left[-\frac{1}{2}\left(\|\boldsymbol{x} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Lambda}_x^{-1}(\sigma)} + \|\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}}\|^2_{\boldsymbol{\Lambda}_\mu^{-1}(\vartheta)} + \|\boldsymbol{\sigma} - \boldsymbol{\mu}_\sigma\|^2_{\boldsymbol{\Lambda}_\sigma^{-1}(\gamma)}\right)\right]. \end{aligned} \tag{19}$$

In practice, $\tilde{\mu}_i$ can be assigned based on the average absorption coefficients of tissue types provided in the literature with a sufficiently large variance ϑ_i^2. Analogously, the mean value of $\boldsymbol{\sigma}$ can be extracted from the error bounds of the average optical properties of tissue types which are documented in the literature. See for example, Mobley and Vo-Dinh (2003) and Cheong *et al* (1990) for an exhaustive list of optical coefficients for human tissue and fluids. An alternative method could be to employ a low-resolution estimate of the optical image (typically a least-squares solution) to extract approximate values for these parameters from the optical data. While alternative distribution models on $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ can be considered, we will see in the next section and appendix that the Gaussian model leads to a computationally efficient hyperparameter estimation scheme.

4. Image reconstruction and hyperparameter estimation

Following an empirical Bayesian approach (Berger 1988), we propose an iterative algorithm to estimate both the optical image and the hyperparameters. At each iteration, the MAP estimates of the hyperparameters are computed by successively maximizing the joint distribution with respect to each hyperparameter. The hyperparameter estimation step at each iteration is followed by an image update.

Substituting equations (9) and (19) into (6), the joint probability distribution of the measurements, optical image and the hyperparameters becomes

$$\begin{aligned} p(\boldsymbol{y}, \boldsymbol{x}, \lambda, \boldsymbol{\mu}, \boldsymbol{\sigma}|C) &= p(\boldsymbol{y}, \lambda|\boldsymbol{x})p(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\sigma}, C)p(\boldsymbol{\mu}|C)p(\boldsymbol{\sigma}|C) \\ &= \frac{1}{K(2\pi)^{3N/2}|\boldsymbol{\Lambda}_\zeta(\lambda)|^{1/2}|\boldsymbol{\Lambda}_x(\boldsymbol{\sigma})|^{1/2}|\boldsymbol{\Lambda}_\mu(\vartheta)|^{1/2}|\boldsymbol{\Lambda}_\sigma(\gamma)|^{1/2}} \\ &\quad \times \exp\left[-\frac{1}{2}\left(\|\boldsymbol{y} - \boldsymbol{W}\boldsymbol{x}\|^2_{\boldsymbol{\Lambda}_\zeta^{-1}(\lambda)} + \|\boldsymbol{x} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Lambda}_x^{-1}(\sigma)}\right.\right. \\ &\quad \left.\left. + \|\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}}\|^2_{\boldsymbol{\Lambda}_\mu^{-1}(\vartheta)} + \|\boldsymbol{\sigma} - \boldsymbol{\mu}_\sigma\|^2_{\boldsymbol{\Lambda}_\sigma^{-1}(\gamma)}\right)\right]. \end{aligned} \tag{20}$$

Let $\Psi(x, \lambda, \mu, \sigma)$ be the objective functional given by

$$\Psi(x, \lambda, \mu, \sigma) = -\log p(y, x, \lambda, \mu, \sigma | C). \tag{21}$$

Then, the image reconstruction becomes an optimization problem in which the objective functional $\Psi(x, \lambda, \mu, \sigma)$ has to be minimized with respect to the image x and the hyperparameters λ, μ and σ.

The hyperparameter estimation problem has been a focus of both the statistical and engineering communities and many procedures have been suggested to date (Mohammad-Djafari 1993, 1996, Zhou *et al* 1994, Molina *et al* 1999, Utsugi 1997). Since the optimization with respect to the optical image itself is a computationally intense problem, it is desirable to keep the computation complexity of hyperparameter estimation to minimum. Therefore, we propose an iterative algorithm based on the empirical Bayesian approach that successively estimates the hyperparameters. The hyperparameter estimation step is followed by the image update by one iteration of CG algorithm, applied with the current estimates of the hyperparameters.

We consider an alternating minimization scheme for the estimation of the mean and standard deviation, where the current estimate of one of the parameters is used to estimate the other. This approach provides a computationally efficient solution to the hyperparameter estimation problem with only $O(N)$ extra operations at each iteration, where N is the number of voxels. The estimation of the noise scale λ is independent of the image model hyperparameters and requires $O(N)$ extra operations when a gradient-based algorithm (such as conjugate gradient algorithm) is used. As a result, the image is updated based on the current estimates of the hyperparameters with $O(NP)$ operations, by one CG iteration.

A pseudocode describing the details of the proposed iterative image and hyperparameter estimation scheme is given in table 1. The outline of the algorithm is as follows: the image estimate is initialized to zero vector at the beginning of the iterations. At each iteration, for the given image update $\hat{x}$, we consider $\hat{x}$ as the hidden variable of the conditional probability $p(y|\lambda)$ and formulate the MAP estimate of the hyperparameter λ, which corresponds to the minimization of the objective functional $\Psi(x, \lambda, \mu, \sigma)$ with respect to λ,

$$\hat{\lambda}_{\mathrm{MAP}} \leftarrow \arg\min_{\lambda} \Psi(\hat{x}, \lambda, \mu, \sigma). \tag{22}$$

Hyperparameters of the image model are estimated in a similar way. We formulate the MAP estimate of the hyperparameter μ, given the image update as the observations and the current estimate of the standard deviation as the hidden variable of the conditional probability $p(x|\mu, C)$. This corresponds to the minimization of the objective functional with respect to μ,

$$\hat{\mu}_{\mathrm{MAP}} \leftarrow \arg\min_{\mu} \Psi(\hat{x}, \lambda, \mu, \hat{\sigma}). \tag{23}$$

Similarly, we formulate the MAP estimate of the hyperparameter σ, given the image update as the observations and the current estimate of the mean as the hidden variable of the conditional probability $p(x|\sigma, C)$. This is equivalent to the minimization of the objective functional with respect to σ:

$$\hat{\sigma}_{\mathrm{MAP}} \leftarrow \arg\min_{\sigma} \Psi(\hat{x}, \lambda, \hat{\mu}, \sigma). \tag{24}$$

Note that the estimate $\hat{\lambda}$ is attained regardless of the value of σ and μ and vice versa (see the appendix). The estimation of the hyperparameters is followed by the update of the image by one CG iteration

$$\hat{x}_{\mathrm{MAP}} \leftarrow \mathrm{CG}_{\mathrm{update}}\{\Psi(x, \hat{\lambda}, \hat{\mu}, \hat{\sigma})\}, \tag{25}$$

Table 1. The modified conjugate gradient algorithm with the proper initializations and embedded hyperparameter estimation sub-routine.

```
begin{Initialize}
        Image: $\hat{x}^{(0)} = 0$
        Hyperparameters: $\hat{\lambda}^{(0)} = 1; \mu = \tilde{\mu}; \hat{\sigma}_i^{(0)} = k_i$ for $i = 1, 2, \ldots, M k_i > 0$
        Gradient vector: $g^{(0)} = -\Omega(\hat{x}^{(0)}) = -\frac{\partial\Psi(x,\lambda,\mu,\sigma)}{\partial x}|_{x=\hat{x}^{(0)}}$
        Search direction: $d^{(0)} = g^{(0)}$
        Termination criterion: $\varepsilon$
        Iteration counter: $n = 0$
end{Initialize}
repeat
begin{Update Image}
          Exact line search: $\alpha^{(n)} = \arg\min_{\alpha>0} \Psi(\hat{x}^{(n)} + \alpha d^{(n)}, \hat{\lambda}^{(n)}, \hat{\mu}^{(n)}, \hat{\sigma}^{(n)})$
          $\hat{x}^{(n+1)} = \hat{x}^{(n)} + \alpha^{(n)} d^{(n)}$
   end{Update Image}
   begin{Estimate Hyperparameters}
          $\hat{\lambda}^{(n+1)} \leftarrow \arg\min_{\lambda} \Psi(\hat{x}^{(n)}, \lambda, \mu^{(n)}, \sigma^{(n)})$
          $\hat{\mu}^{(n+1)} \leftarrow \arg\min_{\mu} \Psi(\hat{x}^{(n)}, \lambda^{(n)}, \mu, \sigma^{(n)})$
          $\hat{\sigma}^{(n+1)} \leftarrow \arg\min_{\sigma} \Psi(\hat{x}^{(n)}, \lambda^{(n)}, \hat{\mu}^{(n+1)}, \sigma)$
   end{Estimate Hyperparameters}
   begin{Update Search Direction}
          $g^{(n+1)} = -\Omega(\hat{x}^{(n+1)})$
          $\beta^{(n+1)} = \max(\frac{g^{(n+1)T}(g^{(n+1)} - g^{(n)})}{g^{(n)T} g^{(n)}}, 0)$
          $d^{(n+1)} = g^{(n+1)} + \beta^{(n+1)} d^{(n)}$
   end{Update Search Direction}
   $n = n + 1$
until $\|\alpha^{(n)} \Omega(\hat{x}^{(n)})\| < \varepsilon$
```

where the step length is computed by the exact line search (Nash and Sofer 1996) and Polak–Ribière–Polyak method (Polak and Ribière 1969, Polyak 1969) is used to calculate β parameter of the CG method. The derivation of the hyperparameters is given in the appendix in detail. The computational complexity of the proposed image reconstruction and hyperparameter estimation algorithm is shown in table 2.

Note that the proposed simultaneous image reconstruction and hyperparameter estimation algorithm can be viewed as a variation of the alternating minimization algorithm (Csiszar and Tusnady 1984), where the minimization of the objective functional $\Psi(x, \lambda, \mu, \sigma)$ with respect to the image is replaced by one CG iteration that leads to the update of the image. Similar approaches can be found in Mohammad-Djafari (1993, 1996) and Milstein *et al* (2002). The empirical approach proposed in this work is asymptotically efficient and comparable with the hierarchical analysis (Berger 1988, Molina *et al* 1999) provided that the number of observations (N_i for the sub-images, P for the measurements) is large (Berger 1988). Alternative approaches for hyperparameter estimation include ML-Type II (Berger 1988) and marginalized ML type estimation schemes, which do not incorporate hyperpriors. Extension of these approaches to marginalized MAP estimation requires integration over the multi-dimensional image x (Molina *et al* 1999, Galatsanos *et al* 2002), which may result in increased computational complexity.

5. Results

We perform three sets of experiments to evaluate the performance of the proposed method. For each of the experiments, we used a finite difference code to simulate the optical measurements.

Table 2. Computation complexity of the algorithm described in table 1. N is the number of voxels and P is the number of measurements. Since the noise scale estimate $\hat{\lambda}$ is used in the calculation of α, this saves $O(NP)$ number of multiplications and $O(NP)$ number of additions. Thus, the estimation sub-routine does not affect the overall computational complexity of the original conjugate gradient algorithm.

Operation	Number of multiplications	Number of additions
α computation (exact line search)	$O(NP)$	$O(NP)$
β parameter calculation	$O(N)$	$O(N)$
Estimation of λ	$O(NP)$	$O(NP)$
Estimation of μ	$O(N)$	$O(N)$
Estimation of σ	$O(N)$	$O(N)$
Gradient calculation and image update	$O(NP)$	$O(NP)$
Total	$O(NP)$	$O(NP)$

3% of the mean value of measurements was added to the measurement vector $\mathbf{y}$ in each experiment.

In the first set of experiments, the objective is to evaluate whether the optical image reconstruction is biased towards the average optical coefficients, which are used to formulate the hyperprior defined on the mean value. Second set of experiments demonstrate how well the hierarchical Bayesian formulation captures the correlation between the optical and the *a priori* anatomical image. In the last set of experiments, we evaluate the proposed method using optical data simulated from an MR breast image. We show that the *a priori* information improves the image reconstruction and does not lead to an erroneous bias towards the *a priori* information.

5.1. Simulation experiment I

A priori selection of parameters in the assumed image and/or noise models may bias the optical image reconstruction in an undesirable way. The hierarchical Bayesian formulation and the empirical hyperparameter estimation scheme proposed in this paper avoids such undesirable results by incorporating dynamic image and noise models in the problem formulation. In this experiment, we show that the proposed hyperparameter estimation approach is relatively insensitive to average optical values used to design the hyperprior $p(\mu|C)$, defined on the mean value of the image.

We consider a square heterogeneity with a mean absorption coefficient of 0.071 cm^{-1} embedded in a background with a mean absorption coefficient of 0.04 cm^{-1} as shown in figure 1. The diffusion coefficient of both the heterogeneity and the background is assumed equal and set to $D = 0.033$ cm. We consider a transmission geometry and distribute 19 sources and 19 detectors on opposite sides to yield a total of 722 measurements, collected at two frequencies, that is 100 and 200 MHz. We evaluate the Jacobian at $\mu_{a0} = 0.04\ \text{cm}^{-1}$ and perform 200 experiments with the proposed hierarchical Bayesian formulation and hyperparameter estimation scheme. At each experiment, $p(\mu|C)$ is formulated such that $\tilde{\mu}_i + \mu_{a0}$ for the square inclusion and for the background are drawn randomly from a uniform distribution with lower and upper bounds (0.038, 0.114) and (0.02, 0.06), respectively (Note that $\tilde{\mu}_i$ value is used in the formulation of the objective functional). The associated standard

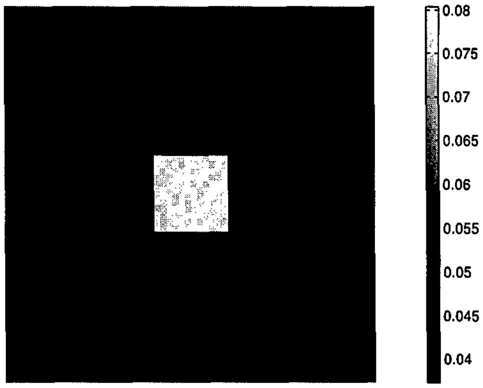
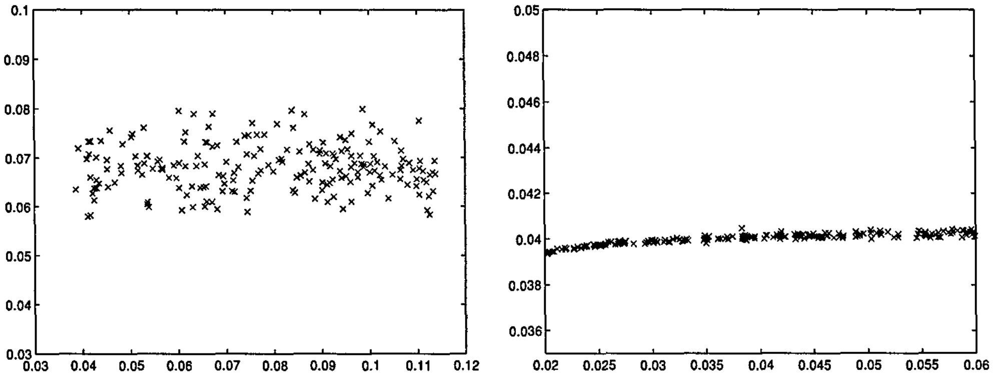


Figure 1. The medium used to simulate the optical data for the first experiment. A square absorber is embedded in an almost homogeneous background.



(a) (b)

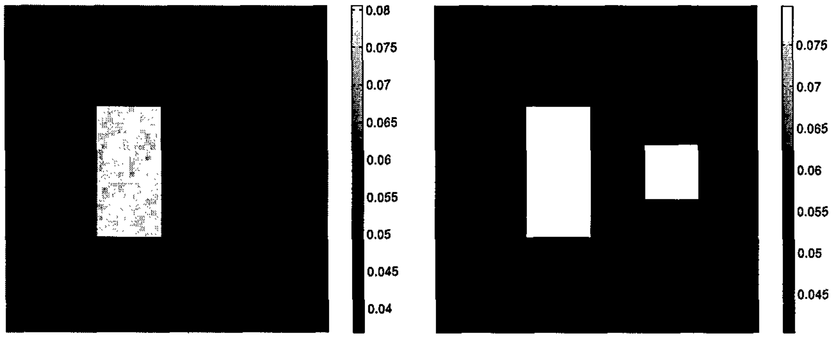
Figure 2. The average absorption values ($\bar{\hat{\mu}}_a$) in the reconstructed sub-images for randomly drawn average absorption values $\tilde{\mu}_a$. Results are shown for 200 trials. Note that the $\bar{\hat{\mu}}_a$ values estimated for the square inclusion (an average value of 0.068) are very close to the actual value of 0.071. (a) The average value $\bar{\hat{\mu}}_a$ in the reconstructed square heterogeneity versus the average absorption value drawn for the square heterogeneity is shown. (b) The average value $\bar{\hat{\mu}}_a$ in the reconstructed background sub-image versus the average absorption value $\tilde{\mu}_a$ drawn for the background is shown.

deviation ϑ_i is sufficiently large and set to $\vartheta_i = 6(\tilde{\mu}_i + \mu_{a0})$. Similarly $p(\sigma|C)$ is formulated such that $\mu_{\sigma_i} = 0.4(\tilde{\mu}_i + \mu_{a0})$ and $\gamma_i = 15\mu_{\sigma_i}$.

This simulation study demonstrates that the proposed method provides effective means to constrain image reconstruction without biasing the solution. Figure 2 shows the average of the estimated absorption values $\bar{\hat{\mu}}_a$ (i.e. $\hat{x}_i + \mu_{a0}$) versus the assigned hyperparameter values $\tilde{\mu}_a$ (i.e. $\tilde{\mu}_i + \mu_{a0}$) for the square inclusion and the background, respectively. We observe that quantitative accuracy is achieved even in the extreme cases and the reconstruction for the background is almost insensitive to the assigned hyperparameters.

5.2. Simulation experiment II

In this experiment, we examine a case where the heterogeneity is present in the *a priori* anatomical image but not in the optical image (see figure 3). Based on the anatomical template, the optical image was segmented into two sub-images, one corresponding to the background and the other corresponding to the two inclusions, which were assumed to have



(a) (b)

Figure 3. The optical image (a) and the anatomical counterpart (b). Note the additional absorber (square inclusion) indicated by the anatomical prior which does not exist in the optical map.

Table 3. The parameter set used in the inverse problem formulations for the simulation experiment II, displayed for the two sub-images: the inclusions and the background. $\mu_{a0} = 0.04$ and $D = 0.33$ cm for this experiment. This experiment was performed with the same source-detector configuration as in experiment 1.

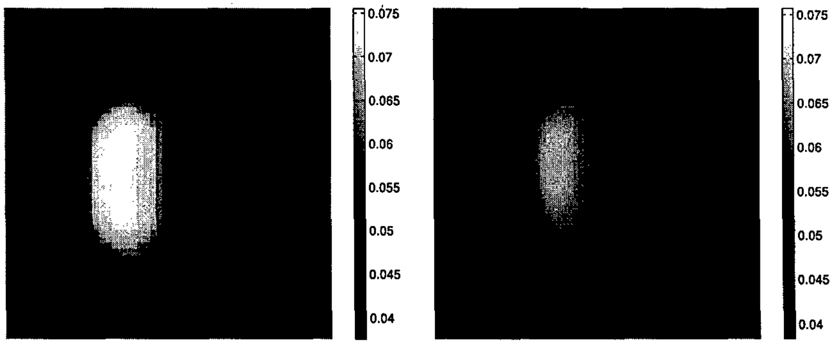
	Hyperprior parameters	
	$(\bar{\mu}_i + \mu_{a0}, \vartheta_i)$	$(\mu_{\sigma_i}, \gamma_i)$
Both inclusions in the anatomical image	(0.08, 0.04)	(0.03, 0.45)
Background	(0.04, 0.02)	(0.02, 0.30)

the same average optical coefficients. For comparison, we also considered the maximum likelihood (ML) approach for the inverse problem formulation. The ML approach estimates the optical image based on the data likelihood model given in section 3.1. This formulation does not incorporate any *a priori* information about the image. Nevertheless, the noise scale λ is unknown and has to be estimated as described in section 4. Note that the ML (no prior) formulation is in fact regularized by the stopping criterion of the conjugate gradient algorithm used in the minimization of the resulting objective functional. We list the parameter set used in the hierarchical prior design in table 3.

The reconstructed images are shown in figure 4. The true mean values and the sample average of the estimated absorption coefficients of the heterogeneity on the left and on the right and of the background are shown in table 4. Even though the anatomical image indicates a heterogeneity on the right, the hierarchical Bayesian formulation leads to a qualitatively good reconstruction. The ML estimate of the image detects the rectangular inclusion, but suffers from low resolution and lacks accuracy in the reconstructed value of the absorption coefficient of the rectangular absorber.

5.3. 2D experiment with MR-simulated data

We used the T1-weighted MR breast image from Ntziachristos *et al* (1999) to design a realistic optical breast model (figure 5). The MR breast image was segmented into parenchyma and adipose layers by applying a simple thresholding algorithm with respect to the MR image intensity values. Next, a tumour corresponding to an infiltrating ductal carcinoma revealed by Gd-DTPA (gadolinium-diethylenetriamine pentaacetic acid) enhancement was inserted (shown in figure 5 as well). Each sub-region was assigned an absorption value as indicated



(a) (b)

Figure 4. The reconstructed images as a result of hierarchical Bayesian formulation (a), and ML (no prior) formulation (b). ML solution is quantitatively inaccurate and suffers from low spatial resolution. Hierarchical Bayesian formulation leads to an image estimate with quantitative accuracy.

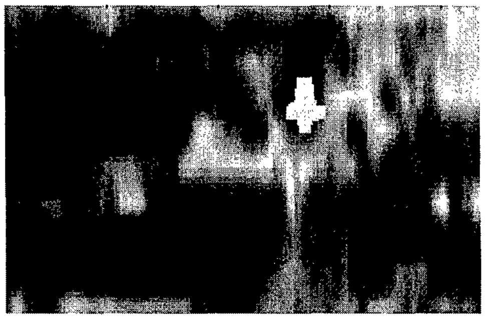
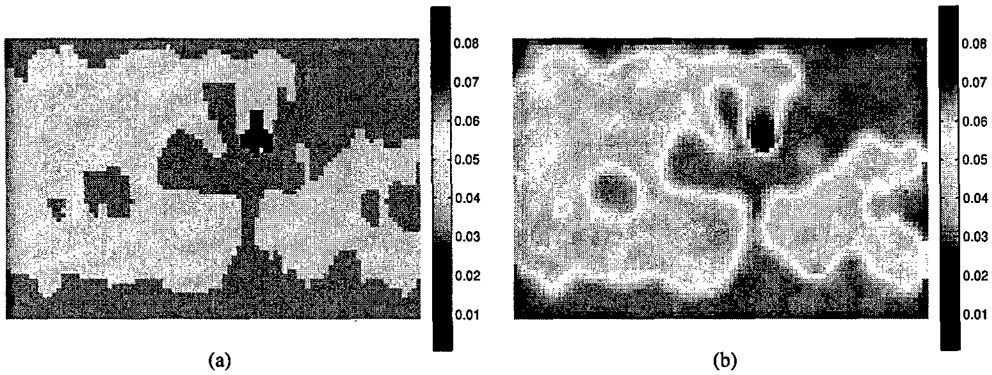


Figure 5. The original MR breast image with an artificial tumour inserted.

Table 4. The actual mean values and the mean of the reconstructed sub-images as a result of maximum likelihood (no prior) and hierarchical Bayesian formulations. Sub-images are defined on the sub-domains as indicated by the anatomical image shown in figure 3(b).

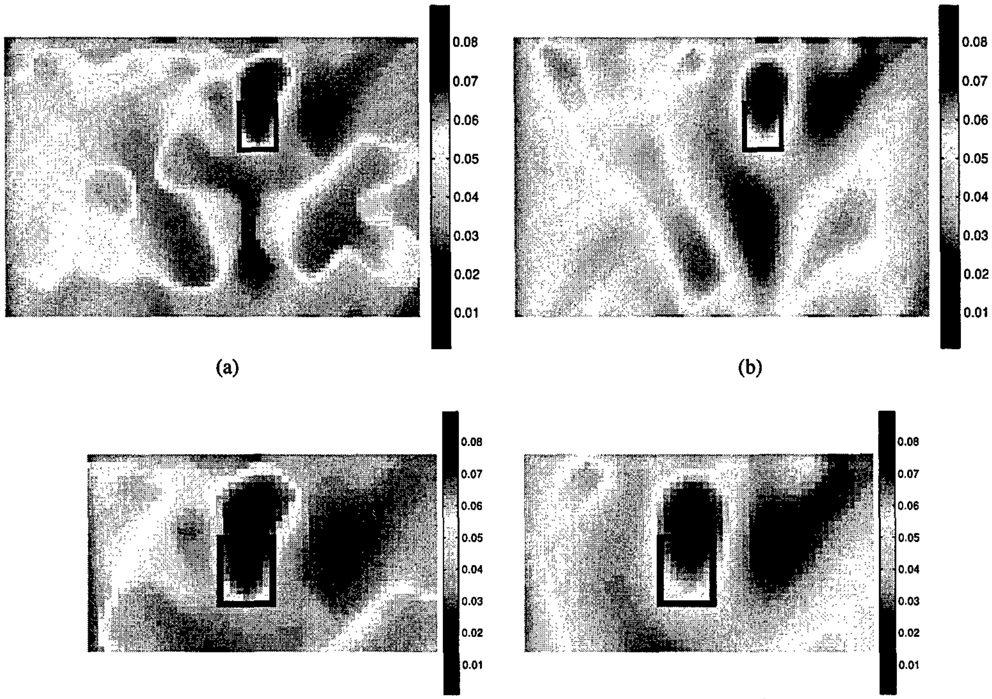
Sub-domain	True mean value	Maximum likelihood	Hierarchical Bayesian
Rectangular inclusion	0.076	0.061	0.071
Square inclusion	0.040	0.037	0.044
Background	0.040	0.041	0.040

in Ntziachristos *et al* (1999) $\left(\mu_a^{\text{adipose}} = 0.03\,\text{cm}^{-1}, \mu_a^{\text{parenchyma}} = 0.06\,\text{cm}^{-1}, \mu_a^{\text{tumour}} = 0.09\,\text{cm}^{-1}\right)$ to obtain an initial template (figure 6(b)). To simulate a corresponding optical image, zero mean Gaussian noise was added prior to filtering the image by a low-pass filter. The resulting optical image is shown in figure 6(b). Note the quantitative and spatial mismatch along the boundaries and especially within the tumour. The homogeneous diffusion coefficient of the medium was set to 0.042 cm. Nine frequencies ranging from 0 to 244 MHz were used to obtain 729 measurements with nine sources and nine detectors positioned along the x-axis on opposite sides. The optical medium was uniformly discretized into 90 pixels along the x-axis and 60 pixels along the y-axis leading to a total of 5400 1×1 cm^2 pixels.


0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01

(a) (b)

Figure 6. The anatomical and optical images are shown on the left and right, respectively. Note the spatial mismatch between the two images.


0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01

(a) (b)

(c) (d)

Figure 7. The hierarchical Bayesian reconstruction of the optical image (a) using the anatomical template shown in figure 6(a) for the design of the hierarchical image model. Part (c) shows the image that zooms into the tumour region in the optical image shown in (a). The ML estimate of the entire image and the sub-image focusing the tumour region are shown in (b) and (d), respectively. The rectangular box in the figures shows the actual location of the tumour.

We performed two types of experiments to test the performance of the proposed hierarchical Bayesian approach for this problem.

(i) *Tumour present both anatomically and optically.* In this experiment, the template extracted from the anatomical image shown in figure 6(a) was used to design the hierarchical image prior. As a result, the optical image was segmented into three sub-images each of which corresponded to the labelled images in the anatomical image as shown in figure 6(a).

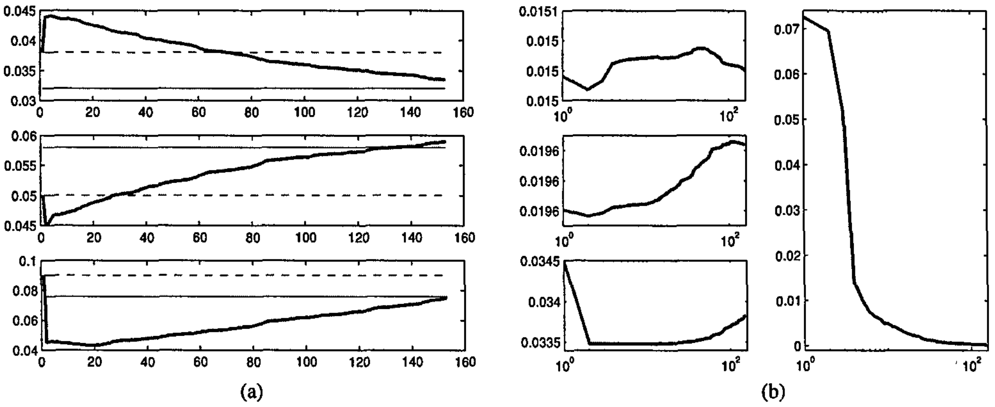


Figure 8. The mean value and the standard deviation estimates for each sub-image (sub-images are determined by the anatomical image) and noise scale estimate versus iteration number are shown. The thick solid line shows the estimated values. The constant solid line in (a) shows the actual mean value and the constant dashed line in (a) shows the assigned mean value ($\bar{\mu}_i + \mu_{a0}$) used in the design of the hyperprior defined on the sub-image means. The actual mean values in these sub-regions are 0.032, 0.058 and 0.076, respectively. (a) The estimated mean values $\hat{\mu}_i$ for each sub-image versus iteration number are shown. The estimates for the parenchyma, adipose and tumour sub-images are given at the top, middle and bottom, respectively. (b) The estimated standard deviation $\hat{\sigma}_i$ for each sub-image versus iteration number are shown on the left. The estimates for the parenchyma, adipose and tumour sub-images are given at the top, middle and bottom, respectively. The noise scale estimate $\hat{\lambda}$ versus iteration number is shown on the right.

In the design of the hyperprior defined on the mean (i.e. $p(\mu|C)$), values that are significantly different from the actual mean of the sub-images were used. Thus, this experiment evaluates the robustness of the proposed method when the true statistics of the optical image are significantly different from the statistics extracted from the prior anatomical image.

The reconstructed image and the sub-image zoomed into the tumour region are shown in figures 7(a) and (c), respectively. For comparison, the ML estimate of the image is shown in figures 7(b) and (d). The simulation results show that hierarchical Bayesian approach leads to qualitatively better results and resolves the tumour more accurately.

In figure 8, the estimates of the hyperparameters associated with the noise and image models are given as a function of the iteration number. Note that the mean value estimates for each sub-image converge to actual values, even though the corresponding assigned hyperparameters regarding the mean value deviate from the true average optical values by at least 15% (see table 5). The experiment also demonstrates that the initialization of the hyperparameters does not have any effect on the performance of the estimation (figure 8).

(ii) *Tumour present optically but not anatomically.* In this experiment, we removed the tumour region from the template extracted from the prior anatomical image, but kept it in the optical image as shown in figures 9(a) and (b). As a result, the optical image was segmented into two sub-images. The objective of this experiment is to evaluate how well the proposed method reconstructs optical tumours when they are not anatomically present.

The reconstructed images for this experiment are given in figures 10(a) and (c), respectively. The ML estimate of the image is given in figures 10(b) and (d). We observe that, even though there is a significant mismatch between the optical image and the anatomical counterpart in the tumour region, the hierarchical Bayesian formulation leads to a qualitatively better reconstruction than the ML approach, even around the tumour. Furthermore, the tumour is better localized as compared to the ML solution and is not biased towards the *a priori*

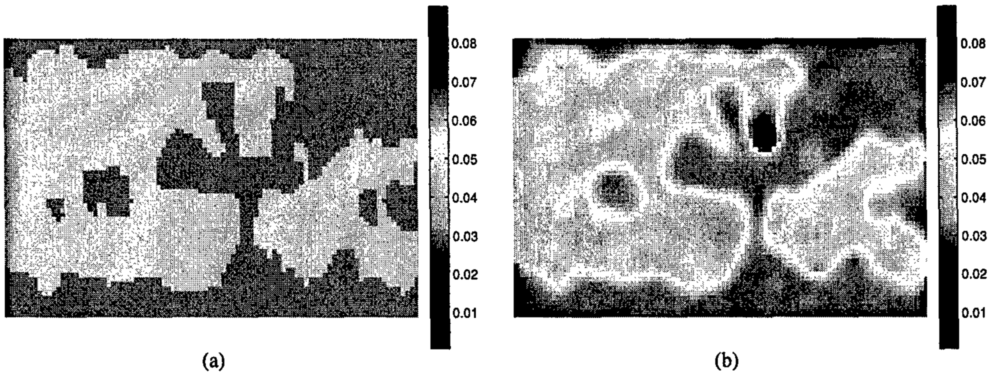


Figure 9. The anatomical template (a) and the original optical image (b). Note that the tumour is not anatomically present.

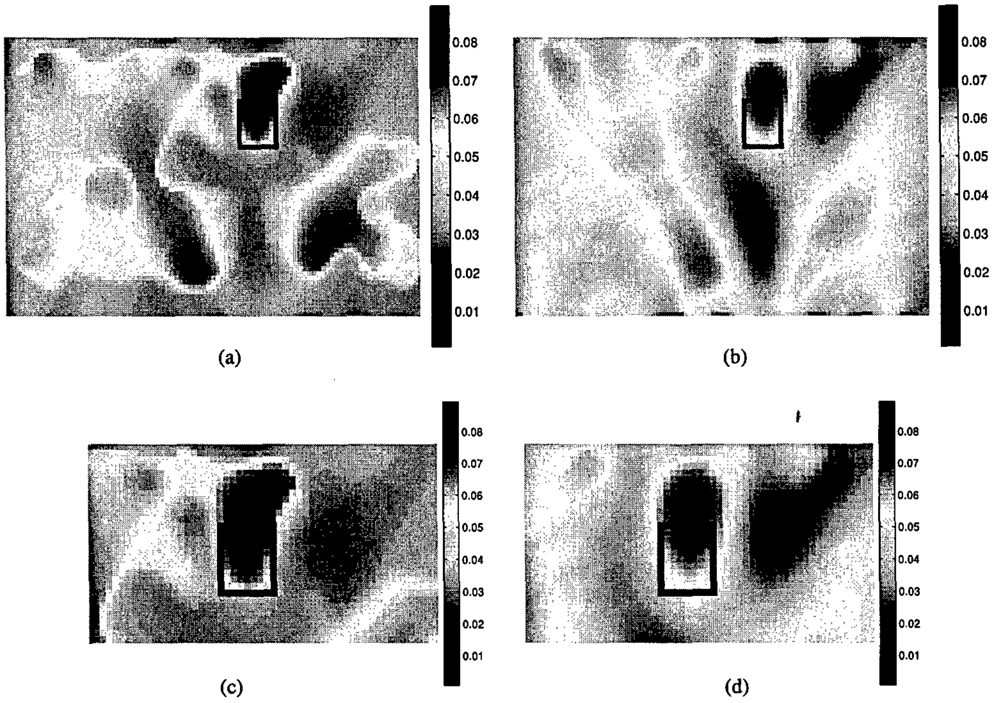
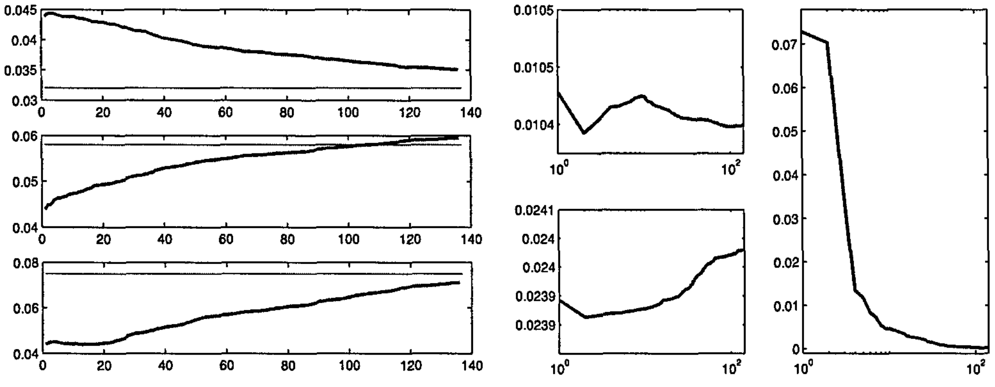


Figure 10. The hierarchical Bayesian reconstruction of the optical image (a) using the anatomical template shown in figure 9(a) for the design of the hierarchical image model. Part (c) shows the image that zooms into the tumour region in the optical image shown in (a). The ML estimate of the entire image and the sub-image focusing the tumour region are shown in (b) and (d), respectively. The rectangular box in the figures shows the actual location of the tumour.

anatomical image. The error in the localization of the tumour can be attributed to the source-detector geometry. The propagation of light along the y-direction results in a smoothing effect on the optical image along the y-direction. This effect is enhanced near source and detectors due to the behaviour of the solution of the diffusion equation. The vertically smoothing effect can be observed in the ML estimate of the image more apparently (figures 10(b) and (d)). The smoothing effect can be suppressed by incorporation of *a priori* information for the tumour



(a) (b)

Figure 11. The mean values of the reconstructed sub-images versus iteration number (a). The estimated values of the standard deviation of the two sub-images and the noise scale λ are shown in (b). (a) The average value $(\bar{\mu}_a$ in each reconstructed sub-image versus iteration number are shown. The estimates for the parenchyma, adipose and tumour sub-images are given at the top, middle and bottom, respectively. The domains of the sub-images correspond to the domains in the optical image shown in figure 9(a). (b) The estimated standard deviation $\hat{\vartheta}_i$ for each sub-image versus iteration number are shown on the left. The estimates for the parenchyma, adipose sub-images are given at the top and bottom, respectively. The sub-images correspond to the sub-images defined by the anatomical template shown in figure 9(b). The noise scale estimate $\hat{\lambda}$ versus iteration number is shown on the right.

Table 5. The actual mean of the absorption values in each sub-image and the parameter set used in the inverse problem formulations for the MR-simulated experiments I and II. $\mu_{a0} = 0.0439$ for this experiment. N/A stands for 'not applicable'.

		Sub-images		
		Parenchyma	Adipose	Tumour
The first experiment	$(\bar{\mu}_i + \mu_{a0}, \vartheta_i)$:	(0.038, 0.23)	(0.05, 0.3)	(0.09, 0.54)
	$(\mu_{\sigma_i}, \gamma_i)$:	(0.015, 0.228)	(0.02, 0.3)	(0.036, 0.54)
The second experiment	$(\bar{\mu}_i + \mu_{a0}, \vartheta_i)$:	(0.03, 0.18)	(0.06, 0.36)	N/A
	$(\mu_{\sigma_i}, \gamma_i)$:	(0.012, 0.18)	(0.024, 0.36)	N/A
	$(\bar{\mu}_a^{\text{actual}} + \mu_{a0})$:	0.032	0.058	0.076

as in case (i), where the tumour is better resolved (figure 7(c)). Further improvement can be achieved by employing sources and detectors positioned along the y-axis as well as along x-axis.

In figure 11, the average absorption value of each reconstructed sub-image versus iteration number is given. The sub-images correspond to those as indicated by the actual optical image shown in figure 6. Note that the mean value of the reconstructed image in the tumour region converges to the actual value even though the anatomical image asserts that no tumour exists.

The set of parameters used in the design of hyperpriors for these experiments and the actual mean of absorption values for each sub-image are shown in table 5.

6. Conclusion

In this work, we formulated the inverse DOT problem within a hierarchical Bayesian framework where the hierarchical prior distribution is based on the *a priori* information extracted from a secondary high-resolution anatomical image. Instead of directly constraining the optical image with the anatomical prior, we incorporated the *a priori* information in the form of hyperpriors to impose constraints on the unknown hyperparameters of the image and noise models. We proposed a computationally efficient iterative algorithm, based on an empirical Bayesian approach, to simultaneously estimate the optical image and the unknown hyperparameters. We tested the proposed approach in three different simulation experiments. Numerical experiments demonstrate that the proposed approach improves the spatial resolution and quantitative accuracy of optical images. Our study shows that the hierarchical Bayesian approach provides an effective framework to capture the correlation between optical and anatomical images.

The proposed hierarchical Bayesian formulation can be extended to incorporate spectral *a priori* information (Intes *et al* 2004). Finally, we note that the results are based on the linearized forward model around a homogeneous background. However, the proposed hierarchical Bayesian formulation and the iterative optical image and hyperparameter estimation scheme can be adapted to the nonlinear inverse DOT problem, wherein the Jacobian of the forward model is iteratively updated.

Acknowledgments

The authors are grateful to M Holboke for developing the finite difference code. Xavier Intes and Britton Chance acknowledge partial support from NIH CA 87046 and RR 02305. Xavier Intes acknowledges partial support from the National Institute of Health grant no CA 110173. Birsen Yazici and Murat Guven acknowledge partial support from NSF-BES-0353160, ONR-N00014-04-1-0694 and US Army Medical Research-W81XWH-04-1-0559.

Appendix

The minimization of the objective function $\Psi(\boldsymbol{x}, \lambda, \boldsymbol{\mu}, \boldsymbol{\sigma})$ with respect to the noise scale λ given the updated image estimate $\hat{\boldsymbol{x}}$ results in the ML estimate $\hat{\lambda}$ of the noise scale:

$$\hat{\lambda} = \frac{1}{P}\|\boldsymbol{y} - \boldsymbol{W}\hat{\boldsymbol{x}}\|^2_{\Lambda_y^{-1}}. \tag{A.1}$$

Note that the estimate $\hat{\lambda}$ is independent of the hyperparameters associated with the image model. On the other hand, the minimization of the objective functional with respect to $\boldsymbol{\mu}$ yields the MAP estimate of the mean. In order to find an estimate for the vector $\boldsymbol{\mu}$, we make use of the probability density function formulation for each individual sub-image and rewrite the objective function with μ_i dependent terms, given the sub-image estimate $\hat{\boldsymbol{x}}_i$ and the current estimate of the standard deviation $\hat{\sigma}_i$:

$$\Psi_{\mu_i}(\boldsymbol{\mu}_i) = \frac{1}{2\hat{\sigma}_i^2}\|\hat{\boldsymbol{x}}_i - \tilde{\boldsymbol{\mu}}_i\|^2 + \frac{1}{2\vartheta_i^2}\|\boldsymbol{\mu}_i - \tilde{\boldsymbol{\mu}}_i\|^2. \tag{A.2}$$

Minimization with respect to yields the estimate $\hat{\mu}_i$ for the mean value of the ith sub-image, i.e.

$$\begin{aligned}\hat{\mu}_i &= \frac{1}{N_i}\left(\frac{\vartheta_i^2}{\vartheta_i^2 + \hat{\sigma}_i^2}\sum_k^{N_i} x_{ik} + \frac{\hat{\sigma}_i^2}{\vartheta_i^2 + \hat{\sigma}_i^2}\sum_k^{N_i} \tilde{\mu}_{ik}\right)\\ &= \frac{1}{N_i}\left(\frac{\vartheta_i^2}{\vartheta_i^2 + \hat{\sigma}_i^2}\sum_k^{N_i} x_{ik} + \frac{\hat{\sigma}_i^2}{\vartheta_i^2 + \hat{\sigma}_i^2} N_i\tilde{\mu}_i\right),\end{aligned} \tag{A.3}$$

where x_{ik} denotes the kth voxel in the ith sub-image and $\tilde{\mu}_{ik} = \tilde{\mu}_i$ stands for the assigned mean value of the image mean for the kth voxel in the ith sub-image.

Following a similar procedure, we make use of the hierarchical prior formulation for each individual sub-image and rewrite the objective functional with σ_i-dependent terms, given the sub-image estimate $\hat{\boldsymbol{x}}_i$ and the current estimate of the mean $\hat{\boldsymbol{\mu}}_i$:

$$\arg\min_{\sigma_i} \Psi_{\sigma_i}(\sigma_i) = \log\left(\sigma_i^{N_i}\right) + \frac{1}{2\sigma_i^2}\|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{\mu}}_i\|^2 + \frac{1}{2\gamma_i^2}\|\boldsymbol{\sigma}_i - \boldsymbol{\mu}_{\sigma_i}\|^2, \qquad \text{(A.4)}$$

where $\boldsymbol{\sigma}_i = (\sigma_i \cdots \sigma_i)^T$. After taking the derivative of the above expression, the estimate for the sub-image standard deviation σ_i satisfies the following equation:

$$N_i\hat{\sigma}_i^2 - \|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{\mu}}_i\|^2 + \frac{N_i}{\gamma_i^2}\hat{\sigma}_i^3(\hat{\sigma}_i - \mu_{\sigma_i}) = 0. \qquad \text{(A.5)}$$

This is a fourth-order equation in $\hat{\sigma}_i$, which is rather difficult to solve. In order to simplify the solution, we make use of the following approximation:

$$\mu_{\sigma_i}\hat{\sigma}_i^3 \approx \mu_{\sigma_i}^2, \qquad \text{(A.6)}$$

and equation (A.5) becomes a quadratic equation of the estimate $\hat{\sigma}_i^2$

$$\frac{N_i}{\gamma_i^2}\hat{\sigma}_i^4 + N_i\hat{\sigma}_i^2 - \left(\|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{\mu}}_i\|^2 + N_i\frac{\mu_{\sigma_i}^2}{\gamma_i^2}\right) = 0. \qquad \text{(A.7)}$$

Then the positive value for the variance estimate $\hat{\sigma}_i^2$ is equal to

$$\hat{\sigma}_i^2 = \frac{-N_i + \sqrt{\Delta}}{2N_i/\gamma_i^2} > 0, \qquad \text{(A.8)}$$

where the discriminant Δ in the equation can be evaluated as

$$\Delta = N_i^2 + 4\frac{N_i}{\gamma_i^2}\left(\|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{\mu}}_i\|^2 + N_i\frac{\mu_{\sigma_i}^2}{\gamma_i^2}\right) > 0. \qquad \text{(A.9)}$$

Note that in the limiting case, for $\gamma_i^2 \to \infty$, the sub-image variance estimate $\hat{\sigma}_i^2$ converges to

$$\lim_{\gamma_i^2\to\infty} \hat{\sigma}_i^2 = \lim_{\gamma_i^2\to\infty} \frac{-N_i + \sqrt{\Delta}}{2N_i/\gamma_i^2} = \frac{\|\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{\mu}}_i\|^2}{N_i}, \qquad i = 1, 2, \ldots, M, \qquad \text{(A.10)}$$

which corresponds to the ML estimate of the variance (Guven *et al* 2003b).

References

Arridge S R 1993 Forward and inverse problems in time resolved infrared imaging *Medical Optical Tomography* vol IS11 ed G Muller (Bellingham, WA: SPIE Optical Engineering) pp 53–64

Arridge S R 1995 Photon measurement density functions part 1: analytical forms *Appl. Opt.* **34** 7395–409

Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–93

Arridge S R and Hebden J C 1997 Optical imaging in medicine: II. Modeling and reconstruction *Phys. Med. Biol.* **42** 841–53

Barbour R L, Graber H L, Chang J, Barbour S-L S, Koo P C and Aronson R 1995 MRI-Guided optical tomography: prospects and computation for a new imaging method *IEEE Comput. Sci. Eng.* **2** 63–76

Berger J O 1988 *Statistical Decision Theory and Bayesian Analysis* (Berlin: Springer)

Boas D, Brooks D, Miller E, DiMarzio C, Kilmer M, Gaudette R J and Zhang Q 2001 Imaging the body with diffuse optical tomography *IEEE Signal Process. Mag.* **18** 57–75

Brooksby B, Dehghani H, Pogue B and Paulsen K D 2003 Near-Infrared (NIR) tomography breast reconstruction with *a priori* structural information from MRI: algorithm development reconstructing heterogeneities *IEEE J. Sel. Top. Quantum Electron.* **9** 199–209

Cerussi A, Berger A, Bevilacqua F, Shah N, Jakubowski D, Butler J, Holcombe R F and Tromberg B J 2001 Sources of absorption and scattering contrast for near-infrared optical mammography *Acad. Radiol.* **8** 211–8

Chance B, Nioka S and Chen Y 2003 Shining new light on brain function *OE Mag.* **3** 16–9

Chang J, Graber H L and Barbour R L 1997 Imaging of fluorescence in highly scattering media *IEEE Trans. Biomed. Eng.* **44** 810–22

Chang J, Graber H L, Koo P C, Aronson R, Barbour S L and Barbour R L 1997 Optical imaging of anatomical maps derived from magnetic resonance imaging using time-independent optical sources *IEEE Trans. Med. Imaging* **16** 68–77

Cheong W F, Prahl S A and Welch A J A 1990 Review of the optical properties of biological tissues *IEEE J. Quantum Electron.* **26** 2166–85

Csiszar I and Tusnady G 1984 Information geometry and alternating minimization procedures *Stat. Decis.* (Suppl) **1** 205–37

Cuccia D J, Bevilacqua F, Durkin A J, Merritt S, Tromberg B J, Gulsen G, Yu H, Wang J and Nalcioglu O 2003 *In vivo* quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration *Appl. Opt.* **42** 2940–50

Dehghani H Pogue B W and Paulsen K D 2002 Development of hybrid NIR/MRI imaging system algorithm: use of *a-priori* information for tumor detection in the female breast *Proc. IEEE Int. Symp. on Biomedical Imaging* pp 657–60

Eppstein M J, Hawrysz D J, Godavarty A and Sevick-Muraca E M 2002 Three dimensional near infrared fluorescence tomography with Bayesian methodologies for image reconstruction from sparse and noisy data sets *Proc. Natl Acad. Sci.* **99** 9619–24

Frangioni J V 2003 *In vivo* near-infrared fluorescence imaging *Curr. Opin. Chem. Biol.* **7** 626–34

Galatsanos N P, Mesarovic V Z, Molina R, Katsaggelos A K and Mateos J 2002 Hyperparameter estimation in image restoration problems with partially-known blurs *Soc. Photo-Opt. Instrum. Eng.* **41** 1845–54

Gulsen G, Birgul O and Nalcioglu O 2003 Hybrid DOT and MR system *Proc. SPIE* **4955** 246–52

Guven M, Yazici B, Intes X and Chance B 2003a Diffuse optical tomography with *a priori* anatomical information *Proc. SPIE* **4955** 634–46

Guven M, Yazici B, Intes X and Chance B 2003b Three-dimensional diffuse optical tomography with *a priori* anatomical information *Proc. SPIE* **5138** 268–80

Guven M, Yazici B, Intes X, Chance B and Zheng Y 2002 Recursive least squares algorithm for optical diffusion tomography *Proc. IEEE 28th Annual Northeast* pp 273–4

Hawrysz D J and Sevick-Muraca E M 2000 Developments toward diagnostic breast cancer imaging using near-infrared optical measurements and fluorescent contrast agents *Neoplasia* **2** 388–417

Hebden J C, Arridge S R and Delpy D T 1997 Optical imaging in medicine: I. Experimental techniques *Phys. Med. Biol.* **42** 825–40

Hebden J C, Gibson A, Austin T, Yusof R, Everdell N, Delpy E, Arridge S R, Meek J H and Wyatt J S 2004 Imaging changes in blood volume and oxygenation in the newborn infant brain using three-dimensional optical tomography *Phys. Med. Biol.* **49** 1117–30

Hielscher A H and Bartel S 2001 Use of penalty terms in gradient-based iterative reconstruction schemes for optical tomography *J. Biomed. Opt.* **6** 183–92

Intes X and Chance B 2005 Non-PET functional imaging techniques: optical *The Radiol. Clin. North Am.* **43** 221–34

Intes X, Maloux C, Guven M, Yazici B and Chance B 2004 Diffuse optical tomography with physiological and spatial *a priori* constraints *Phys. Med. Biol.* **49** N155–63

Intes X, Ripoll J, Chen Y, Nioka S, Yodh A and Chance B 2003 *In vivo* continuous-wave optical breast imaging enhanced with Indocyanine Green *Med. Phys. Biol.* **30** 1039–47

Intes X, Yu J, Yodh A G and Chance B 2002 Development and evaluation of a multi-wavelength multi-channel time resolved optical instrument for NIR/MRI mammography co-registration *Proc. IEEE 28th Annual Northeast Bioengineering Conf. (IEEE* Cat. No.02CH37342*)* pp 91–2

Jiang H, Paulsen K D, Osterberg U L, Pogue B W and Patterson M S 1996 Optical image reconstruction using frequency-domain data: simulation and experiment *J. Opt. Soc. Am.* A **13** 253–66

Kak A C and Slaney M 1988 *Principles of Computerized Tomographic Imaging* (New York: IEEE) pp 208–18

Kincade K 2004 Optical diagnostics continue migration from bench top to beside *Laser Focus World* **40** 130–4

Li A *et al* 2003 A tomographic optical breast imaging guided by 3D mammography *Appl. Opt.* **42** 5181–90

Merritt S, Bevilacqua F, Durkin A J, Cuccia D J, Lanning R, Tromberg B J, Gulsen G, Yu G, Wang J and Nalcioglu O 2003a Coregistration of diffuse optical spectroscopy and magnetic resonance imaging in a rat tumor model *Appl. Opt.* **42** 2951–58

Merritt S, Gulsen G, Chiou G, Chu Y, Deng C, Cerussi A, Durkin A, Tromberg B and Nalcioglu O 2003b Comparison of water and lipid content measurements using diffuse optical spectroscopy and MRI in emulsion phantoms *Technol. Cancer Res. Treat.* **2** 563–9

Milstein A B, Oh S, Reynolds J S, Webb K J and Bouman C A 2002 Three-dimensional Bayesian optical diffusion tomography with experimental data *Opt. Lett.* **27** 95–7

Mobley J and Vo-Dinh T 2003 *Biomedical Photonics, Handbook* (Boca Raton, FL: CRC Press) chapter 2

Mohammad-Djafari A 1993 On the estimation of hyperparameters in Bayesian approach of solving inverse problems *IEEE Int. Conf. ICASSP Acoustics, Speech, and Signal Processing* vol 5 pp 495–8

Mohammad-Djafari A 1996 Joint estimation of parameters in a Bayesian approach of solving inverse problems *Proc. IEEE Int. Conf. on Image Processing (Lausanne, Switzerland)* vol 2 pp 473–76

Molina R, Katsaggelos A K and Mateos J 1999 Bayesian and regularization methods for hyperparameter estimation in image restoration *IEEE Trans. Image Process.* **8** 31–46

Nash S G and Sofer A 1996 A linear and nonlinear programming *McGraw-Hill Industrial Engineering Series*

Ntziachristos V, Chance B and Yodh A G 1999 Differential diffuse optical tomography *Opt. Exp.* **5** 230–42

Ntziachristos V, Yodh A G, Schnall M and Chance B 2000 Concurrent MRI and diffuse optical tomography of breast after Indocyanine Green enhancement *Proc. Natl. Acad. Sci. USA* **97** 2767–80

Ntziachristos V, Yodh A G, Schnall M and Chance B 2002 MRI-guided diffuse optical spectroscopy of malignant and benign breast legions *Neoplasia* **4** 347–54

Oh S, Milstein A B, Millane R P, Bouman C A and Webb K J 2002 Source-detector calibration in three-dimensional Bayesian optical diffusion tomography *J. Opt. Soc. Am.* A **19** 1983–93

O'Leary M A 1996 Imaging with diffuse photon density waves *PhD Thesis* University of Pennsylvania

O'Leary M A, Boas D A, Chance B and Yodh A G 1995 Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography *Opt. Lett.* **20** 426–28

Paulsen K D and Jiang H 1996 Enhanced frequency domain optical image reconstruction in tissues through total variation minimization *Appl. Opt.* **35** 3447–58

Pogue B W, McBride T O, Prewitt J, Osterberg U L and Paulsen K D 1999 Spatially varying regularization improves diffuse optical tomography *Appl. Opt.* **38** 2950–61

Pogue B W and Paulsen K D 1998 High-resolution near-infrared tomographic imaging simulations of the rat cranium by use of *a priori* magnetic resonance imaging structural information *Opt. Lett.* **23** 1716–18

Polak E and Ribière G 1969 Note sur la convergence de directions conjugèes *Rev. Francaise Informat Recherche Opertionelle, 3 eannèe* **16** 35–43

Polyak B T 1969 The conjugate gradient method in extremem problems *USSR Comput. Math. Phys.* **9** 94–112

Quaresima V, Lepanto R and Ferrari M 2003 The use of near infrared spectroscopy in sports medicine *J. Sports Med. Phys. Fit.* **43** 1–13

Schnall M D 2003 Breast MR imaging *Radiol. Clin. North Am.* **41** 43–50

Schweiger M and Arridge S R 1999 Optical tomographic reconstruction in a complex head model using *a priori* region boundary information *Phys. Med. Biol.* **44** 2703–21

Srinivasan S, Pogue B, Jiang S, Dehghani H, Kogel C, Soho S, Gibson J J, Tosteson T D, Poplack S P and Paulsen K D 2003 Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured *in vivo* by near-infrared breast tomography *Proc. Natl Acad. Sci.* **100** 12349–54

Strangman G, Boas D and Sutton J 2002 Non-invasive neuroimaging using near-infrared light *Biol. Psychiatry* **52** 679–93

Utsugi A 1997 Hyperparameter selection for self-organizing maps *Neural Comput.* **9** 623–35

Villringer A and Chance B 1997 Non-invasive optical spectroscopy and imaging of human brain function *Trends Neurosci.* **20** 435–42

Weissleder R and Ntziachristos V 2003 Shedding light onto live molecular targets *Nat. Med.* **9** 123–8

Xu H, Dehghani H, Pogue B W, Paulsen K D and Dunn J D 2002 Hybrid MR/Near infrared imaging of the muribe brain: optimization of optical fiber arrangement and use of *a priori* knowledge *Proc. IEEE Int. Symp. on Biomedical Imaging* pp 74–7

Yao Y, Wang Y, Pei Y, Zhu W and Barbour R L 1997 Frequency domain optical imaging of absorption and scattering distributions by a Born iterative method *J. Opt. Soc. Am.* A **14** 325–42

Ye J C, Bouman C A, Webb K J and Millane R P 2001 Nonlinear multigrid algorithms for Bayesian optical diffusion tomography *IEEE Trans. Image Process.* **10** 909–22

Yodh A G and Chance B 1995 Spectroscopy and imaging with diffusing light *Phys. Today* **48** 34–40

Zhou Z, Leahy R M and Mumcuoglu E U 1994 Simultaneous hyperparameter estimation and Bayesian image reconstruction for PET 1994 *IEEE Conference Record Nuclear Science Symp. and Medical Imaging Conf.* vol 4 pp 1604–08

Zhu Q, Chen N and Kurtzman S H 2003a Imaging tumor angiogenesis by use of combined near-infrared diffusive light and ultrasound *Opt. Lett.* **28** 337–9

Zhu Q, Huang M, Chen N, Zarfos K, Jagjivan B, Kane M, Hedge P and Kurtzman S H 2003b Ultrasound-guided optical tomographic imaging of malignant and benign breast lesions: initial clinical results of 19 cases *Neoplasia* **5** 379–88

INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **49** (2004) N155–N163 PII: S0031-9155(04)75972-2

NOTE

Diffuse optical tomography with physiological and spatial *a priori* constraints

Xavier Intes[1,2,4], Clemence Maloux[1], Murat Guven[3], Birzen Yazici[3] and Britton Chance[1]

[1] Department of Biophysics and Biochemistry, University of Pennsylvania, Philadelphia, PA 19104, USA
[2] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
[3] Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

E-mail: xintes@art.ca

Received 5 February 2004
Published 26 May 2004
Online at stacks.iop.org/PMB/49/N155
DOI: 10.1088/0031-9155/49/12/N01

Abstract

Diffuse optical tomography is a typical inverse problem plagued by ill-condition. To overcome this drawback, regularization or constraining techniques are incorporated in the inverse formulation. In this work, we investigate the enhancement in recovering functional parameters by using physiological and spatial *a priori* constraints. More accurate recovery of the two main functional parameters that are the blood volume and the relative saturation is demonstrated through simulations by using our method compared to actual techniques.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Probing human tissue with near infrared (NIR) light is emerging as a new, promising imaging modality. The strength of this new technology relies on its ability to reveal the functional state of deep tissue. To date the main applications of NIR technologies are functional brain imaging (Strangman *et al* 2002, Villringer and Chance 1997), muscle imaging (Lin *et al* 2002) and optical mammography (Colak *et al* 1999, Franceschini *et al* 1997, Grosenick *et al* 2003, Intes *et al* 2003, Jiang *et al* 2002, McBride *et al* 2001, Tromberg *et al* 2000). The last application is known to be of growing interest in both the research and medical community (Cerussi and Tromberg 2003).

Optical mammography aims to retrieve important, local physiological parameters that are the blood content and the relative oxygenation of the blood. These two functional

---

[4] Now with ART Advanced Research Technologies.

0031-9155/04/120155+09$30.00 © 2004 IOP Publishing Ltd Printed in the UK

parameters are known to provide a means of discriminating between healthy and diseased tissues. The blood content relates to the angiogenesis level of the tumourous mass, and the relative oxygenation to its hypermetabolic state. These two functional signatures correlate to malignancy.

Pre-clinical data strengthen these assumptions. However, difficulties arise due to the specific nature of the light propagation. In this spectral window, the light is strongly diffused leading to a relatively poor resolution for thick tissue investigation. Also, the breast is by nature a heterogeneous organ with a complex spatial distribution of optical relevant chromophores. In the NIR spectral range, four chromophores are accountable for the absorption (Cerussi *et al* 2001). These chromophores are the oxy- ([HbO_2]), deoxy-haemoglobin ([Hb]), the water ([H_2O]) and the lipids ([Li]). The first two chromophores are providing an insight into the tissue functional state. The last two chromophores are linked to the structural architecture of the breast. The spatial distribution and the relative concentrations of these chromophores are patient dependent and even more can vary over time in the same patient due to hormonal regulation (Chance 2001, Cubeddu *et al* 2000, Durduran *et al* 2002, Shah *et al* 2001, Srinivasan *et al* 2003).

To enhance diffuse optical tomography (DOT) performances, researchers have proposed fusing optical techniques with other medical imaging modality. Magnetic resonance imaging (MRI) is the perfect candidate for optical co-registration (Brooksby *et al* 2003, Guven *et al* 2004, Ntziachristos *et al* 2002, Pei *et al* 1999, Pogue and Paulsen 1998). MRI provides high spatial resolution maps of the breast optical structure that are relevant to the water and lipid distribution. Moreover, MRI can provide a means to estimate the concentration of these two structural chromophores (Merrit *et al* 2003). In this paper, we investigate the first step towards incorporating physiological and spatial *a priori* information derived from MRI.

2. Methods

2.1. Forward model

The propagation of NIR light in tissue is well modelled by the diffusion equation. In the case of heterogeneity, the diffusion equation can be solved by a perturbative approach (O'Leary 1996). In this work, we have used the Rytov approximation approach. In the case of DOT, multiple source–detector pairs are used. The medium under consideration is sampled in voxels and the problem can be written as a matrix equation (Arridge 1999), i.e.:

$$\begin{pmatrix} \Phi(r_{s1}, r_{d1}) \\ \vdots \\ \Phi(r_{sm}, r_{dm}) \end{pmatrix} = \begin{pmatrix} W_{11} & \dots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{m1} & \dots & W_{mn} \end{pmatrix} \times \begin{pmatrix} \delta\mu_a(r_1) \\ \vdots \\ \delta\mu_a(r_n) \end{pmatrix} \tag{1}$$

where $\Phi(r_{si}, r_{di})$ is the diffuse perturbative phase for the *i*th source–detector pair, W_{ij} (O'Leary 1996) is the weight for the *j*th voxel and the *i*th source–detector pair and $\delta\mu_a(r_j)$ is the differential absorption coefficient of the *j*th voxel. We limited our problem to image the absorption coefficient. Boundary conditions for semi-infinite geometries and slab geometries are derived using the extrapolated boundary condition and the image source technique (Haskell *et al* 1994).

2.2. Functional imaging

The estimation of the absorption coefficient at several wavelengths enables the provision of spatial maps of the targeted chromophores. In the case of the breast, the four chromophores of potential diagnostic interest are the oxy- and deoxy-haemoglobin, the water and the lipid.

The concentrations of these breast constituents are linearly related to the absorption values through the linear system:

$$\begin{pmatrix} \delta\mu_a^{\lambda 1}(r_j) \\ \vdots \\ \delta\mu_a^{\lambda p}(r_j) \end{pmatrix} = \begin{pmatrix} \varepsilon_{\mathrm{Hb}}^{\lambda 1} & \varepsilon_{\mathrm{HbO_2}}^{\lambda 1} & \varepsilon_{\mathrm{H_2O}}^{\lambda 1} & \varepsilon_{\mathrm{Li}}^{\lambda 1} \\ \vdots & \vdots & \vdots & \vdots \\ \varepsilon_{\mathrm{Hb}}^{\lambda p} & \varepsilon_{\mathrm{HbO_2}}^{\lambda p} & \varepsilon_{\mathrm{H_2O}}^{\lambda p} & \varepsilon_{\mathrm{Li}}^{\lambda p} \end{pmatrix} \times \begin{pmatrix} \delta[\mathrm{Hb}](r_j) \\ \delta[\mathrm{HbO_2}](r_j) \\ \delta[\mathrm{H_2O}](r_j) \\ \delta[\mathrm{Li}](r_j) \end{pmatrix} \tag{2}$$

where $\delta\mu_a^{\lambda k}(r_j)$ is the differential absorption coefficient at the kth wavelength λ_k and for the jth voxel, $\varepsilon_C^{\lambda k}$ is the extinction coefficient of the Cth chromophore at the kth wavelength and $\delta[C](r_j)$ is the differential concentration of the Cth chromophore for the jth voxel.

Solving this linear system on a voxel basis for a spectral set of absorption distributions gives the functional maps required for diagnostic purposes. We will refer to this approach as 'indirect imaging' through this paper.

Recently, to overcome the burden and minimize the systematic errors due to the ill-condition of both inverse problems, a method aiming at directly imaging the functional parameters has been proposed (Durduran *et al* 2001). This method takes advantage of the linear relationship described previously and formulates the inverse problem as

$$\begin{pmatrix} \Phi^{\lambda 1}(r_{s1}, r_{d1}) \\ \vdots \\ \Phi^{\lambda 1}(r_{sm}, r_{dm}) \\ \vdots \\ \Phi^{\lambda p}(r_{s1}, r_{d1}) \\ \vdots \\ \Phi^{\lambda p}(r_{sm}, r_{dm}) \end{pmatrix} = \begin{pmatrix} \varepsilon_{\mathrm{Hb}}^{\lambda 1} W^{\lambda 1} & \varepsilon_{\mathrm{HbO_2}}^{\lambda 1} W^{\lambda 1} & \varepsilon_{\mathrm{H_2O}}^{\lambda 1} W^{\lambda 1} & \varepsilon_{\mathrm{Li}}^{\lambda 1} W^{\lambda 1} \\ \varepsilon_{\mathrm{Hb}}^{\lambda 2} W^{\lambda 2} & \varepsilon_{\mathrm{HbO_2}}^{\lambda 2} W^{\lambda 2} & \varepsilon_{\mathrm{H_2O}}^{\lambda 2} W^{\lambda 2} & \varepsilon_{\mathrm{Li}}^{\lambda 2} W^{\lambda 2} \\ \vdots & \vdots & \vdots & \vdots \\ \varepsilon_{\mathrm{Hb}}^{\lambda p} W^{\lambda p} & \varepsilon_{\mathrm{HbO_2}}^{\lambda p} W^{\lambda p} & \varepsilon_{\mathrm{H_2O}}^{\lambda p} W^{\lambda p} & \varepsilon_{\mathrm{Li}}^{\lambda p} W^{\lambda p} \end{pmatrix} \times \begin{pmatrix} \delta[\mathrm{Hb}](r_j) \\ \delta[\mathrm{HbO_2}](r_j) \\ \delta[\mathrm{H_2O}](r_j) \\ \delta[\mathrm{Li}](r_j) \end{pmatrix}. \tag{3}$$

This new linear system is poorly conditioned and great care should be taken during the pre-conditioning of this sensitivity matrix. In our case, we used an average column scheme (Pei *et al* 2001) applied to each sub-matrix of the kernel as written in (3). This specific pre-conditioning scheme led to the most accurate reconstructions. We will refer to the scheme expressed in (3) as 'direct imaging' through this document.

The functional parameters that we are interested in are the blood volume:

$$[\mathrm{BV}] = [\mathrm{Hb}] + [\mathrm{HbO_2}] \tag{4}$$

and the relative saturation:

$$[\mathrm{SaO_2}] = \frac{[\mathrm{HbO_2}]}{[\mathrm{Hb}] + [\mathrm{HbO_2}]}. \tag{5}$$

2.3. *Bayesian framework*

Due to the ill-posed and/or underdetermined nature of the DOT problem, the solution of the inverse problem is not typically robust. One avenue to overcome this difficulty is to incorporate *a priori* information constraining the space of unknowns.

The Bayesian approach provides a natural framework to incorporate prior information. Guven *et al* (2004) proposed an algorithm based on the Bayesian framework with a spatially varying *a priori* probability density function extracted from MRI anatomical maps. Here we propose to extend this algorithm to a spatial physiological prior. The full derivation of the theoretical developments for the algorithm is presented in detail in Guven *et al* (2004) and we follow herein the same mathematical expressions. We present in this investigation only the salient features of the theoretical approach that are relevant to this work.

The available high-resolution anatomical image is segmented into sub-images that represent major tissue types (typically: parenchyma, glandular and tumour). Prior probability density function of the image is formulated in such a way that each sub-image is assigned a mean value that need not be equal to its actual optical value; and a 'confidence level' is defined in the form of an image variance formulation to allow local variations within sub-images. As a consequence, the overall formulation of the prior information becomes spatially varying, which is specific to the image of interest. Maximum *a posteriori* (MAP) estimate of the image is formed based on the formulation of the image's probability density function

$$\hat{x}_{\mathrm{MAP}} = \arg\max_{x}\{\log p(y|x) + \log p(x)\} \qquad (6)$$

where $p(y|x)$ is the data likelihood function and $p(x)$ is the probability density function of the corresponding image. An 'alternating minimization' algorithm, which sequentially updates the unknown parameters, is used to solve the resulting optimization problem.

For our purposes, the probability density function of the ith sub-image, as defined in the spatial prior, becomes

$$p(\boldsymbol{x}_i|\sigma_i) = \frac{1}{\left(2\pi\sigma_i^2\right)^{N_i/2}} \exp\left(-\frac{1}{2\sigma_i^2}\|\boldsymbol{x}_i - \boldsymbol{C}_i\|^2\right) \qquad i = 1, 2, \ldots, M \quad (7)$$

where M is the number of sub-regions and N_i is the number of voxels in the ith sub-image, $\boldsymbol{x}_i$ is the unknown sub-image, $\boldsymbol{C}_i$ is the assigned chromophore mean concentration and σ_i^2 the single variance, which is an unknown parameter (estimated during the solution). To incorporate the confidence level into the statistical reconstruction procedure, the sub-image variances are expressed as

$$p(\sigma_i) = \frac{1}{\left(2\pi\gamma_i^2\right)^{N_i/2}} \exp\left(-\frac{1}{2\gamma_i^2}\|\sigma_i - \bar{\sigma}_i\|^2\right) \qquad i = 1, 2, \ldots, M \qquad (8)$$

where γ_i is the variance and $\bar{\sigma}_i$ the mean value of σ_i. These two last parameters are *a priori* defined by the user. Thus the confidence level incorporated into the statistical reconstruction procedure is defined on the chromophores concentration. Hence, physiological priors are implicitly defined and allow constraining the reconstruction in physiologically meaningful ranges.

2.4. Measurement generation

Measurements were obtained by solving the frequency-domain diffusion equation with a finite difference approach (FDM). We restricted our simulations to a two-dimensional (2D) geometry for computational efficiency and to a continuous wave data set type. The slab thickness was 6 cm simulating a soft compressed breast. We placed nine sources on one side of the slab and nine detectors on the other side, both evenly stretched along 8 cm (cf figure 1). Two square inclusions of 1 cm^2 were simulated in this model.

The functional properties were chosen to mimic typical values encountered in the human breast. Table 1 summarizes the functional parameters simulated.

The optical models were computed for six wavelengths to replicate the spectral information gathered by our time resolved instrument that performs co-registration with MRI (Intes *et al* 2002). The subsequent optical properties are compiled in table 2. The FDM computations were performed with a 1 mm mesh resolution. The sources and detectors were positioned 2 cm away from the edges to avoid any boundary effects.

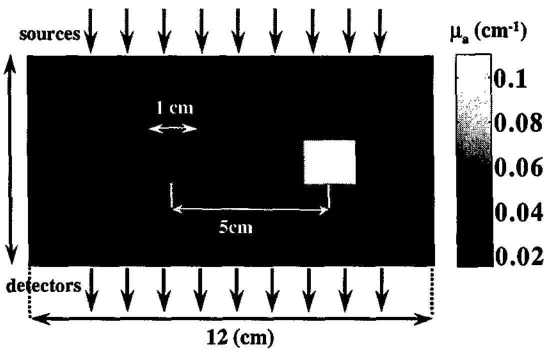


Figure 1. Optical model used for the simulations. The optical properties displayed here correspond to $\lambda = 750$ nm.

Table 1. Functional parameters used to define the optical properties to generate the synthetic measurements.

	Background	Left object	Right object
BV (mM)	0.02	0.06	0.12
SaO_2 (%)	70	65	50

Table 2. Optical parameters for the spectral set investigated herein. No [H_2O] and [Li] constituents were used in these simulations for simplicity.

	690	750	780	805	830	850
μ_s' (cm^{-1})	9.268	8.386	8.000	7.646	7.425	7.216
$\mu_a^{\mathrm{bkg.}}$ (cm^{-1})	0.016	0.016	0.016	0.016	0.018	0.019
μ_a^{Left} (cm^{-1})	0.054	0.050	0.050	0.049	0.053	0.056
μ_a^{Right} (cm^{-1})	0.140	0.115	0.107	0.095	0.100	0.105

3. Results

First, the estimation of the functional parameters simulated is accomplished with a conjugate gradient descent (CGD) algorithm in the case of indirect imaging without using any kind of prior. The results are displayed in figure 2.

Then reconstructions using the indirect imaging approach within the Bayesian framework are provided in figure 3. First, an estimation of the absorption coefficients at the six wavelengths was performed using spatial *a priori* information and a conjugate gradient algorithm with the Polak–Ribiere method (Polak and Ribiere 1969). The absorption coefficients mean values assigned were the simulated ones with a 30% level of confidence (Guven *et al* 2004). Then classical spectroscopy was performed on the resulting optical maps using the linear system of (3) reduced to [HbO_2] and [Hb] concentrations. Last, the results using physiological and spatial *a priori* information are depicted in figure 4. The chromophore concentration means assigned were the exact ones with a 30% level of confidence.

The functional quantitative values retrieved from each case and for the relevant region of interest are summarized in table 3.

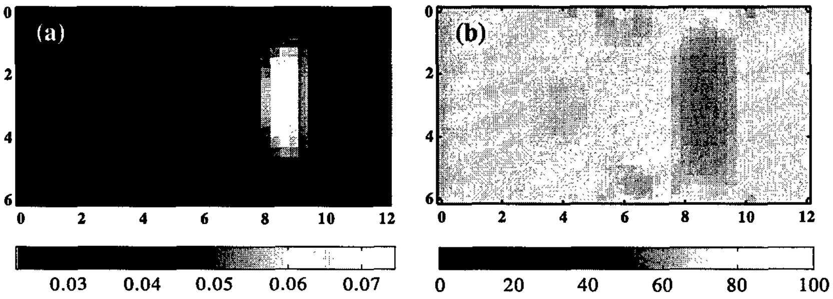


Figure 2. (a) Blood volume (mM) reconstructions and (b) saturation estimates (%) in the case of classical indirect imaging.

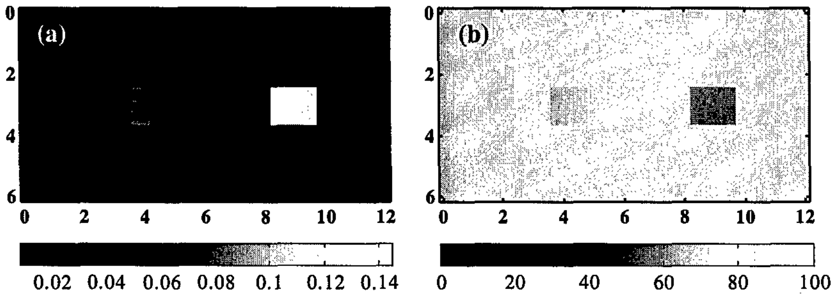


Figure 3. (a) Blood volume (mM) reconstructions and (b) saturation estimates (%) in the case of Bayesian indirect imaging. The assigned mean values of the absorption correspond to the true absorption values with a 30% level of confidence.

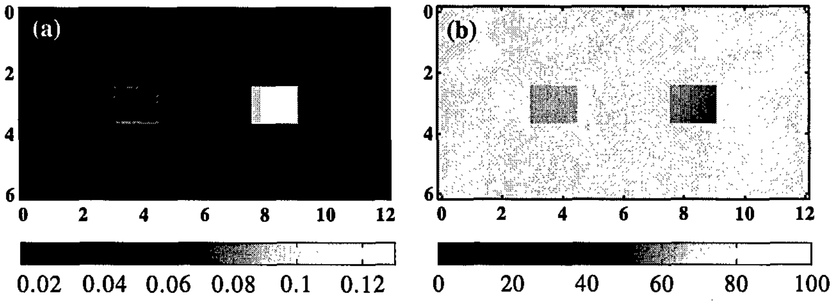


Figure 4. (a) Blood volume (mM) reconstructions and (b) saturation estimates (%) in the case of Bayesian direct imaging. The assigned mean values correspond to the true concentration values with a 30% level of confidence.

4. Discussion

In all the cases investigated herein, the two inclusions were reconstructed. As expected, the incorporation of spatial *a priori* information increased the image resolution. In the case of indirect imaging without *a priori* information the two inclusions are elongated and thus the contrast is diluted. This fact is reflected in the estimation of the blood volume that is underestimated for both objects (~35% lower estimate). When *a priori* anatomical information is incorporated within the Bayesian framework, the estimation of the blood volume is more

Table 3. Functional parameters recovered with the three approaches and for the three different functional areas. The values proposed here correspond to the mean value of the entire ROI defined as the *a priori* spatial masks. [BV] in mM and [SaO_2] in %.

	Background		Left object		Right object	
	[BV]	[SaO_2]	[BV]	[SaO_2]	[BV]	[SaO_2]
Indirect imaging	0.023	71	0.042	67	0.075	57
Bayesian indirect imaging	0.019	72.46	0.065	63.73	0.122	55.44
Bayesian direct imaging	0.020	69.95	0.061	65.27	0.127	50.02

accurate falling to less than 10% of misestimation in the worse case (1% in the best case). Moreover, when direct imaging is applied in this framework, the two objects are recovered respectively within 1% and 5% for the left and right objects compared to 8% and 2% in the case of indirect imaging.

The enhancement of the technique is even more notable when SaO_2 is considered. The recovery of SaO_2 is then more robust but also less prone to artefacts. Especially in the case of the right object, the recovery of the SaO_2 is superior to the direct imaging approach within the Bayesian framework.

Overall, these simulations highlight the potential of our approach to provide more accurate maps of the relevant functional diagnostic parameters. These simulations were limited to [BV] and SaO_2 recovery. The approach is easily extensible to [H_2O] and [Li]. Especially, the ability of MRI to provide spatial concentration of these chromophores for the assigned prior strengthens our approach (Merrit *et al* 2003).

The mean values of the chromophores and the anatomical *a priori* information are defined by the user. The anatomical maps are provided by the structural MRI maps (such as T1) and tumour delineation could be performed by Gd enhanced MRI. Moreover, estimation of the water, lipid and BV concentrations are feasible with MRI. These values can be used for the mean concentration priors. Also, an alternative is to obtain these values with a simple algorithm such as diffuse optical spectroscopy (Ntziachristos *et al* 2002). The paradigm of uniform concentrations for a certain tissue type is overcome in the Bayesian formulation by defining a level of confidence. These effects of confidence ensure the reconstructions are in a physiologically meaningful range and allow the recovery of heterogeneous concentration in a certain kind of tissue type. Such heterogeneous estimates are visible in figures 3 and 4.

This last point is of paramount importance in the case of MRI-assisted DOT. The contrast provided by MRI is not expected to be exactly congruent with the optical contrast. Allowing heterogeneous bounded reconstructions within each sub-image compensates for discrepancies between both modalities.

5. Conclusion

We reported in this work our first step towards incorporating physiological and spatial *a priori* information derived from MRI to assist DOT. Better estimates of the main functional parameters that are [BV] and [SaO_2] were achieved. More accurate functional maps can lead to an increase of the sensitivity and specificity of optical techniques by incorporating structural contrasts that are known to occur in breast cancer. Moreover, these accurate estimations will be important for application such as therapy monitoring (Zhang *et al* 2003).

This preliminary work will be continued with experimental validation and incorporated in our ongoing clinical trial at the University of Pennsylvania. Also, the investigation will be extended to the potential of our approach to recover exogenous contrast agent concentrations. New classes of contrast agents that are tumour specific (Licha 2002) and/or compatible with both technologies (Josephon *et al* 2002) will be investigated. The inclusion of spatial and temporal *a priori* information is expected to provide a better estimate of the fluorochrome concentration but also to probe physiological parameters such as extravasation (Intes *et al* 2003, Cuccia *et al* 2003) that are still elusive with DOT for thick tissue investigation.

Acknowledgments

XI and BC acknowledge partial support from the National Institute of Health grant no CA 87046.

References

Arridge S 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–R93

Brooksby B, Dehghani H, Pogue B and Paulsen K 2003 Near-infrared tomography breast image reconstruction with *a priori* structural information from MRI: algorithm development for reconstructing heterogeneities *IEEE J. Sel. Top. Quantum Electron.* **9** 199–209

Cerussi A, Berger A, Bevilacqua F, Shah N, Jakubowski D, Butler J, Holcombe R and Tromberg B 2001 Sources of absorption and scattering contrast for near-infrared optical mammography *Acad. Radiol.* **8** 211–8

Cerussi A and Tromberg B 2003 Optical mammography inches closer to the clinics *Biophotonics* 38–42

Chance B 2001 Near-infrared optical spectroscopy characterizes breast tissue hormonal and age status *Acad. Radiol.* **8** 209–10

Colak S, van der Mark M, Hooft G, Hoogenraad J, van der Linden E and Kuijpers F 1999 Clinical optical tomography and NIR spectroscopy for breast cancer detection *IEEE J. Sel. Top. Quantum Electron.* **5** 1143–58

Cubeddu R, D'Andrea C, Pifferi A, Taroni P, Torricelli A and Valentini G 2000 Effects of the menstrual cycle on the red and near-infrared optical properties of the human breast *Photochem. Photobiol.* **72–73** 383–91

Cuccia D, Bevilacqua F, Durkin A, Merritt S, Tromberg B, Gulsen G, Yu H, Wang J and Nalcioglu O 2003 *In vivo* quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration *Appl. Opt.* **42** 2940–50

Durduran T, Choe R, Culver J, Zubkov L, Holboke M, Giammarco J, Chance B and Yodh A G 2002 Bulk optical properties of healthy female breast tissue *Phys. Med. Biol.* **47** 2847–61

Durduran T, Giammarco J, Culver J, Choe R, Zubkov L, Holboke M, Intes X, Nioka S, Chance B and Yodh A 2001 Utilizing *a priori* spectral knowledge in diffuse optical tomography *United Engineering Foundation Conf. Advances in Optics for Biotechnology, Medicine and Surgery VII*

Franceschini M, Moesta K, Fantini S, Gaida G, Gratton E, Jess H, Mantulin W, Seeber M, Schlag P and Kaschke M 1997 Frequency-domain techniques enhance optical mammography: initial clinical results *Proc. Natl Acad. Sci. USA* 6468–73

Grosenick D, Moesta T, Wabnitz H, Mucke J, Stroszcynski C, Macdonald R, Schlag P and Rinnerberg H 2003 Time-domain optical mammography: initial clinical results on detection and characterization of breast tumors *Appl. Opt.* **42** 3170–86

Guven M, Yazici B, Intes X, Maloux C and Chance B 2004 Three-dimensional diffuse optical tomography with *a priori* anatomical information *J. Biomed. Opt.* submitted

Haskell R, Svaasand L, Tsay T T, Feng Tc, McAdams M and Tromberg B 1994 Boundary conditions for the diffusion equation in radiative transfer *J. Opt. Soc. Am.* A **11** 2727–41

Intes X, Yu J, Yodh A G and Chance B 2002 Development and evaluation of a multi-wavelength multi-channel time resolved optical instrument for NIR/MRI mammography co-registration *Proc. IEEE 28th Annual Northeast Bioengineering Conf. (IEEE Cat. No.02CH37342)* pp 91–2

Intes X, Ripoll J, Chen Y, Nioka S, Yodh A and Chance B 2003 *In vivo* continuous-wave optical breast imaging enhanced with Indocyanine Green *Med. Phys.* **30** 1039–47

Jiang H, Iftimia N, Eggert J, Fajardo L and Klove K 2002 Near-infrared optical imaging of the breast with model-based reconstruction *Acad. Radiol.* **9** 186–94

Josephon L, Kircher M, Mahmood U, Tang Y and Weissleder R 2002 NIR fluorescent nanoparticles as combined MR/optical imaging probes *Bioconjug. Chem.* **13** 554–60

Licha K 2002 Contrast agents for optical imaging *Top. Curr. Chem.* **222** 1–29

Lin Y, Lech G, Nioka S, Intes X and Chance B 2002 Noninvasive, low-noise, fast imaging of blood volume and deoxygenation changes in muscles using light-emitting diode continuous-wave imager *Rev. Sci. Instrum.* **73** 3065–74

McBride T, Pogue B, Jiang S, Osterberg U and Paulsen K 2001 Initial studies of *in-vivo* absorbing and scattering heterogeneity in near-infrared tomographic breast imaging *Opt. Lett.* **26** 822–4

Merritt S, Gulsen G, Chiou G, Chu Y, Deng C, Cerussi A, Durkin A, Tromberg B and Nalcioglu O 2003 Comparison of water and lipid content measurements using diffuse optical spectroscopy and MRI in emulsion phantoms *Technol. Cancer Res. Treat.* **2** 563–9

Ntziachristos V, Yodh A, Schnall M and Chance B 2002 MRI-guided diffuse optical spectroscopy of malignant and benign breast legions *Neoplasia* **4** 347–54

O'Leary M 1996 Imaging with diffuse photon density waves *PhD Thesis* University of Pennsylvania

Pei Y, Bao F and Barbour R 1999 Modeling of sensitivity and resolution to an included object in homogeneous scattering media and in MRI-derived breast maps *Opt. Exp.* **5** 203–19

Pei Y, Graber H and Barbour R 2001 Normalized constraint algorithm for minimizing inter-parameter crosstalk in DC optical tomography *Opt. Exp.* **9** 97–109

Pogue B and Paulsen K 1998 High-resolution near-infrared tomographic imaging simulations of the rat cranium by use of *a priori* magnetic resonance imaging structural information *Opt. Lett.* **23** 1716–8

Polak E and Ribiere G 1969 Note sur la convergence de directions conjuguees *Rev. Fr. Inform. Rech. O.* **16** 35–43

Shah N, Cerussi A, Eker B, Espinoza J, Butler J, Fiskhin J, Hornung R and Tromberg B 2001 Non-invasive functional optical spectroscopy of human breast tissue *Proc. Natl Acad. Sci. USA* **98** 4420–5

Srinivasan S, Pogue B, Jiang S, Dehghani H, Kogel C, Soho S, Gibson J, Tosteson T, Poplack S and Paulsen K 2003 Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured *in vivo* by near-infrared breast tomography *Proc. Natl Acad. Sci. USA* **100** 12349–54

Strangman G, Boas D and Sutton J 2002 Non-invasive neuroimaging using near-infrared light *Biol. Psychiatry* **52** 679–93

Tromberg B, Shah N, Lanning R, Cerussi A, Espinoza J, Pham T, Svaasand L and Butler J 2000 Non-invasive *in vivo* characterization of breast tumors using photon migration spectroscopy *Neoplasia* **2** 26–40

Villringer A and Chance B 1997 Non-invasive optical spectroscopy and imaging of human brain function *Trends Neurosci.* **20** 435–42

Zhang J, Tu T, Sunar U, Intes X, Nioka S, Zhang Z, Kilger A, Lustig R, Loevner L and Chance B 2003 Application of I&Q detection system in scouting the curative effect of neck squamous cell carcinoma *Proc. SPIE Optical Tomography and Spectroscopy of Tissue V* **4955** 575–80

Extended Kalman Filtering for the Modeling and Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods

Burak Alacam[1], *Student Member, IEEE*, *Birsen Yazıcı[1], *Member, IEEE*
Xavier Intes[2], *Member, IEEE*, and Britton Chance[3], *Member, IEEE*

[1]Department of Electrical, Computer and System Engineering,
Rensselaer Polytechnic Institute 110 8th Street, Troy, NY 12180
Phone: (518) 276-2905, Fax: (518) 276-6261
E-mail: yazici@ecse.rpi.edu
[2]Advanced Research Technologies Inc.
2300 Alfred-Nobel Blvd., Saint-Laurent, Quebec, H4S 2A4 CANADA
[3]Department of Biochemistry and Biophysics,
University of Pennsylvania, Philadelphia, PA 19104

This work was supported by U.S. Army Medical Research Acquisition Activity under grant W81XWH-04-1-0559, and Office of Naval Research under grant N000014-04-1-0694. XI acknowledges partial support from the National Institute of Health grant No: CA110173.

*Corresponding author

Abstract

Compartmental modeling of indocyanine green (ICG) pharmacokinetics, as measured by near infrared (NIR) techniques, has the potential to provide diagnostic information for tumor differentiation. In this paper, we present three different compartmental models to model the pharmacokinetics of ICG in cancerous tumors. We introduce a systematic and robust approach to model and analyze ICG pharmacokinetics based on the extended Kalman filtering (EKF) framework. The proposed EKF framework effectively models multiple-compartment and multiple-measurement systems in the presence of measurement noise and uncertainties in model dynamics. It provides simultaneous estimation of pharmacokinetic parameters and ICG concentrations in each compartment. Moreover, the recursive nature of the Kalman filter estimator potentially allows real time monitoring of time varying pharmacokinetic rates and concentration changes in different compartments. Additionally, we introduce an information theoretic criteria for the best compartmental model order selection, and residual analysis for the statistical validation of the estimates. We tested our approach using the ICG concentration data acquired from four Fischer rats carrying adenocarcinoma tumor cells. Our study indicates that, in addition to the pharmacokinetic rates, EKF model may provide parameters that may be useful for tumor differentiation.

Index Terms

Extended Kalman filter, indocyanine green, compartmental analysis, pharmacokinetics, tumor characterization.

I. INTRODUCTION

Near infrared (NIR) diffuse optical imaging and spectroscopy methods provide quantitative functional information that cannot be obtained by the conventional radiological methods [1], [2], [3]. NIR techniques can provide *in vivo* measurements of the oxygenation and vascularization states, uptake and release of optical contrast agents, and chromophore concentrations with high sensitivity. In particular, NIR diffuse optical techniques in conjunction with optical contrast agents have the potential to characterize angiogenesis, and to differentiate between malignant and benign tumors [4], [5], [6], [7].

At present, indocyanine green (ICG) is the only NIR optical agent approved for human use. In NIR measurements, the presence of ICG within an imaging volume results in an increased signal that can be observed over the course of the experiment. Study of the time kinetics of ICG concentration curves may provide physiologically relevant information for tumor differentiation. Specifically, cancerous tissue types are expected to show high and fast uptake due to the proliferation of "leaky" angiogenetic microvessels, while normal and fatty tissue show little uptake.

A number of research groups reported compartmental modeling of ICG time-kinetic measurements using NIR methods for tumor diagnosis in animal and human subjects [8], [9], [10]. Compartmental model is a mathematical description of the concentrations of contrast agents in which each compartment represents kinetically distinct tissue type. It consists of a set of coupled ordinary partial differential equations (ODE) and a measurement model. Coefficients of the ODE's are the physiological parameters of interest that represent rates of exchange between different compartments. These parameters are non-linearly related to the total concentration of ICG measured by NIR methods. Furthermore, concentration of ICG in each compartment cannot be directly measured non invasively by NIR techniques, making the pharmacokinetic parameter estimation a highly non-linear problem.

Current methods of ICG compartmental modeling involve curve fitting methods and various techniques for solving differential equations. Gurfinkel et al. presented a two-compartment model for ICG kinetics and estimated model parameters [8]. The measurements were obtained using a frequency domain photon migration system coupled with a charge-coupled device. The pharmacokinetic parameters are estimated for each pixel based on a curve fitting method. This study indicates that model parameters show no difference in the ICG uptake rates between normal and diseased tissue. Cuccia et al. presented a study of the dynamics of ICG in an adenocarcinoma rat tumor model [9]. A two-compartmental model describing the ICG dynamics is used to quantify physiologic parameters related to capillary permeability. The ICG concentration curves were fitted to the compartmental model using a non-linear least squares Levenberg-Marquart algorithm. It was shown that different tumor types have different capillary permeability rates. Intes et al. presented the uptake of ICG by breast tumors using a continuous wave diffuse optical tomography apparatus [10]. A two-compartment model is used to analyze the pharmacokinetics of ICG. A curve fitting algorithm, namely the non-linear Nelder-Mead simplex search, is used to estimate the pharmacokinetic parameters. This study shows that the malignant cases exhibit slower rate constants (uptake and outflow) as compared to healthy tissue.

While the studies described above demonstrate the feasibility of the ICG pharmacokinetics in tumor characterization; due to highly non-linear nature of the pharmacokinetic parameter estimation, variation in parameter values from one subject to another, and sparse data available in clinical and laboratory settings, a *systematic and robust* approach is needed to model, estimate and analyze ICG pharmacokinetics. Such an approach must include: *i*) a method for compartmental model order selection, *ii*) a robust method of estimating ICG pharmacokinetic parameters, and *iii*) a method of validating the selected model and the estimation results.

In this paper, we present three different compartmental models for the ICG pharmacokinetics in cancerous tumors and propose an extended Kalman filtering (EKF) framework to estimate the model parameters. The models capture the transportation of ICG between the vascular and extravascular compartments, including interstitial fluid region, parenchymal cell, intracellular binding site, and extravascular, extracellular spaces (EES). Kalman filter (KF) is an optimal recursive modeling and estimation method with numerous advantages in ICG pharmacokinetic modeling. These include: *i*) effective modeling of multiple compartments, and multiple measurement systems governed by coupled ordinary differential equations, in the presence of measurement noise and uncertainties in the compartmental model dynamics; *ii*) simultaneous estimation of pharmacokinetic model parameters and ICG concentrations in each compartment, which is not accessible *in vivo* by means of NIR techniques; *iii*) recursive estimation of time-varying pharmacokinetic model parameters; *iv*) statistical validation of estimated concentrations and error bounds on the pharmacokinetic parameter estimates; *v*) incorporation of available a priori information about the initial conditions of the permeability rates into the estimation procedure; *vi*) potential real-time monitoring of ICG pharmacokinetic parameters and ICG concentrations in different compartments due to the recursive nature of EKF estimation method. Additionally, we present a method of selecting the optimal compartmental model order based on Bayesian information criteria, and a statistical validation method based on the residual analysis.

We tested our approach using the ICG concentration data acquired from four Fisher rats carrying adenocarcinoma tumor cells. Two-, three- and four-compartment models are fitted to data and pharmacokinetic model parameters and concentrations in different compartments are estimated using the EKF framework. Bayesian information criteria suggests that the two-compartment model provides a sufficient fit for our data. The estimated model order and the model parameters are further validated by the residual analysis. The model parameters were used to differentiate between two types of cancerous tumors. Our study suggests that the permeability rates out of the vasculature are higher in edematous tumors as compared to necrotic tumors. Additionally, we observed that in the two-compartment model, the ICG concentration curve is higher in the EES compartment in edematous tumors. This suggests that the ratio of the peak value of the ICG concentrations in different compartments may be a useful parameter to differentiate tumors.

The paper is organized as follows: In Section II, we present the two-, three- and four-compartment models for the ICG pharmacokinetics in tissue. In Section III, we present the state-space representation of the compartmental models; estimation of ICG pharmacokinetics parameters and ICG concentrations in the EKF framework; and optimal model order selection criteria. In Section IV, we present the experimental results obtained from Fischer rat data. Section V summarizes our results and conclusion. The appendix includes the derivation of the likelihood function used in the Bayesian information criteria.

II. ICG PHARMACOKINETIC MODELING USING NIR MEASUREMENTS

A. Indocyanine Green

ICG is an optical dye commonly used in retinopathy and hepatic diagnostics. Given its low toxicity and FDA approval, it has recently been utilized as a blood pooling agent for the detection and diagnosis of cancerous tumors by means of NIR optical methods. The absorption peak of ICG is 805 nm and the fluorescence peak is at 830 nm. ICG has strong affinity for blood proteins. In plasma, ICG is near-completely bound, primarily to albumin. As a result, its *in vivo* kinetics are similar to those of a 70 kD molecule, although it has a molecular weight of about 700 D [11], [12], [13], [14], [15].

ICG is eliminated from the body primarily through the bile. Outside of the circulatory system, it is not available for removal until it returns to the system. The kinetics of this transition offers a potential means of non-invasively assessing the leakiness of large molecules from the microvasculature; this permeability is a characteristic of the poorly developed vasculature observed in angiogenesis. The increase in local microvasculature density is also expected to induce increased perturbation in the optical signal from intercapillary ICG.

There are some differences in the delivery of ICG between normal and cancerous vasculature. In normal tissue, ICG acts as a blood flow indicator in tight capillaries of normal vessels. However in tumor, ICG may act as a diffusible (extravascular) flow in the leaky capillary of cancer vessels. Additionally, the permeability rate is expected to increase as the malignancy advances [9], [10]. Figure 1 (a) and (b) illustrates the ICG flow for healthy and malignant tissue, respectively.

B. Compartmental Analysis of ICG Pharmacokinetics

Compartmental modeling allows relatively simple and effective mathematical representation of complex biological responses due to contrast agents. A region of interest is assumed to consist of a number of compartments, generally representing a volume or a group of similar tissues into which the contrast agent is distributed. The concentration change in a specific compartment is modeled as a result of the exchange of contrast agent between connected compartments. These changes are modeled by a collection of coupled ODEs; each equation describing the time change dictated by the biological laws that govern the concentration exchanges between the interacting compartments [16], [17], [18], [19]. In this work, we investigate three different compartmental models for the ICG kinetics and determine the optimal model order, based on Bayesian information criteria.

1) The four-compartment model: Figure 2 illustrates the capillary and extracapillary space relevant to the four compartment model. The four-compartment model includes capillary region, interstitial fluid region, parenchymal cell region and intracellular binding site as compartments [20]. The ICG, injected intravenously into the subject, can pass through from capillary into reversible binding site inside the cell through the interstitial fluid region and the parenchymal cell region [20], [21], [22]. Moreover, in advanced tumor stages, the leakiness around the tumor vessels is expected to increase, resulting in higher permeability rates during the transportation of ICG into the compartments. A block diagram of the four-compartment transport and chemical model of ICG delivery is shown in Figure 3(a).

Let C_p, C_i, C_{pc}, C_b denote the ICG concentrations in plasma, interstitial fluid region, parenchymal cell region and intracellular binding site, respectively; and let $k_{out}^{(4)}$, $k_a^{(4)}$, $k_b^{(4)}$, $k_c^{(4)}$, $k_d^{(4)}$, $k_e^{(4)}$ and $k_f^{(4)}$ be the constants used as equilibrium coefficients as shown in Figure 3(a). Then the set of differential equations representing the ICG transition between the four compartments are given as follows:

The leakage into and the drainage out of plasma:

$$\frac{dC_p(t)}{dt} = k_b^{(4)} C_i(t) - k_a^{(4)} C_p(t) - k_{out}^{(4)} C_p(t). \tag{1}$$

The leakage into and the drainage out of interstitial fluid region:

$$\frac{dC_i(t)}{dt} = k_a^{(4)} C_p(t) - k_b^{(4)} C_i(t) - k_c^{(4)} C_i(t) + k_d^{(4)} C_{pc}(t). \tag{2}$$

The leakage into and the drainage out of parenchymal cell:

$$\frac{dC_{pc}(t)}{dt} = k_c^{(4)} C_i(t) - k_d^{(4)} C_{pc}(t) - k_e^{(4)} C_{pc}(t) + k_f^{(4)} C_b(t). \tag{3}$$

The leakage into and the drainage out of intracellular binding site:

$$\frac{dC_b(t)}{dt} = k_e^{(4)} C_{pc}(t) - k_f^{(4)} C_b(t). \tag{4}$$

Physiologically, the equilibrium constants are defined by the permeability surface area products given as $PS\rho$, where P is the capillary permeability constant, S is the capillary surface area, and ρ is the tissue density. $k_{out}^{(4)}$ is proportional to the flow rate into and out of the capillary and $k_a^{(4)}$, $k_b^{(4)}$, $k_c^{(4)}$, $k_d^{(4)}$, $k_e^{(4)}$, and $k_f^{(4)}$ represent intra-tissue physiologic effects during ICG delivery from capillary to binding site. Note that the superscript denotes the order of the compartmental model.

The actual bulk ICG concentration in the tissue measured by NIR spectroscopy, $m(t)$, is a linear combination of the ICG concentrations in four different compartments.

$$m(t) = v_p^{(4)} C_p(t) + v_i^{(4)} C_i(t) + v_{pc}^{(4)} C_{pc}(t) + v_b^{(4)} C_b(t), \tag{5}$$

where $v_p^{(4)}$, $v_i^{(4)}$, $v_{pc}^{(4)}$, $v_b^{(4)}$, are volume fractions of plasma, interstitial fluid region, parenchymal cell region and intracellular binding site, respectively.

2) The three-compartment model: In this model, the parenchymal cell and intracellular binding site compartments are combined to form a single compartment called parenchymal cell. This amounts to the assumptions that the transition of ICG into the intracellular binding site is negligible as compared to the other compartments, and therefore omitted from the model. A block diagram of the three-compartment transport and chemical model of ICG delivery is shown in Figure 3(b). Three-compartment transport equations are given as follows:

The leakage into and the drainage out of plasma:

$$\frac{dC_p(t)}{dt} = k_b^{(3)} C_i(t) - k_a^{(3)} C_p(t) - k_{out}^{(3)} C_p(t) \tag{6}$$

The leakage into and the drainage out of interstitial space:

$$\frac{dC_i(t)}{dt} = k_a^{(3)} C_p(t) - k_b^{(3)} C_i(t) - k_c^{(3)} C_i(t) + k_d^{(3)} C_{pc}(t) \tag{7}$$

The leakage into and the drainage out of parenchymal cell:

$$\frac{dC_{pc}(t)}{dt} = k_c^{(3)} C_i(t) - k_d^{(3)} C_{pc}(t) \tag{8}$$

The total ICG concentration measured by NIR:

$$m(t) = v_p^{(3)} C_p(t) + v_i^{(3)} C_i(t) + v_{pc}^{(3)} C_{pc}(t) \tag{9}$$

where $v_p^{(3)}$, $v_i^{(3)}$, $v_{pc}^{(3)}$ and C_p, C_i, C_{pc} are as defined in the four-compartment model.

3) The two-compartment model: In the two-compartment model, the tumor region is assumed to be composed of two compartments, namely the plasma and EES [9], [23], [24]. EES is defined as the region that lies outside of both the vascular region and the tumor cells. The transition of the ICG to the third and fourth compartments are assumed to be negligible. Therefore the last two compartments in the four compartment model is omitted. We consider transcapillary leakage to occur only at the tumor site. We also assume that a small perturbation of the global plasma concentration does not affect the bulk removal. Figure 3(c) shows the block diagram of the two-compartment model for the ICG kinetics. Let C_p and C_e denote the ICG concentrations in the plasma and EES, respectively. Then the two-compartment ICG chemical transport equations are given as follows:

The leakage into and the drainage out of plasma:

$$\frac{dC_p(t)}{dt} = k_b^{(2)} C_e(t) - k_a^{(2)} C_p(t) - k_{out}^{(2)} C_p(t). \tag{10}$$

The leakage into and the drainage out of the EES:

$$\frac{dC_e(t)}{dt} = k_a^{(2)} C_p(t) - k_b^{(2)} C_e(t). \tag{11}$$

The parameters $k_a^{(2)}$ and $k_b^{(2)}$ govern the leakage into and the drainage out of the EES, respectively. The parameter $k_{out}^{(2)}$ describes the ICG elimination from the body through kidneys and liver.

Actual bulk ICG concentration in the tissue measured by NIR is a linear combination of plasma and EES ICG concentrations given by:

$$m(t) = v_p^{(2)} C_p(t) + v_e^{(2)} C_e(t), \tag{12}$$

where the parameters $v_p^{(2)}$ and $v_e^{(2)}$ denote the plasma and EES volume fractions, respectively.

III. EXTENDED KALMAN FILTERING FOR THE ICG PHARMACOKINETICS

For the rest of our discussion, we shall use the explicit form of the two-compartment model as a running example to clarify our notation.

A. State-space Representation of the ICG Pharmacokinetics

Coupled differential equations resulting from the two-compartment modeling of the ICG pharmacokinetics can be expressed in state-space representation as follows:

$$\begin{bmatrix} \dot{C}_e(t) \\ \dot{C}_p(t) \end{bmatrix} = \begin{bmatrix} -k_b^{(2)} & k_a^{(2)} \\ k_b^{(2)} & -(k_a^{(2)} + k_{out}^{(2)}) \end{bmatrix} \begin{bmatrix} C_e(t) \\ C_p(t) \end{bmatrix} + \omega(t), \tag{13}$$

$$m(t) = \begin{bmatrix} v_e^{(2)} & v_2^{(2)} \end{bmatrix} \begin{bmatrix} C_e(t) \\ C_p(t) \end{bmatrix} + \eta(t)$$

where $\omega(t)$ and $\eta(t)$ are uncorrelated zero mean Gaussian processes with covariance matrix Q, and variance σ^2, respectively.

The closed form of the continuous time state-space representation for the $n-$compartment model is given by:

$$d\mathbf{C}(t) = \kappa(\boldsymbol{\alpha}_n)\mathbf{C}(t)dt + \boldsymbol{\omega}(t)dt,$$

$$m(t) = V(\boldsymbol{\alpha}_n)\mathbf{C}(t) + \eta(t). \tag{14}$$

In equation (14), $\mathbf{C}(t)$ denotes the concentration vector; $\kappa(\boldsymbol{\alpha}_n)$ is the KF system matrix, $V(\boldsymbol{\alpha}_n)$ is the KF measurement matrix and $\boldsymbol{\alpha}_n$ is the parameter vector whose elements are the pharmacokinetic constants and the volume fractions for the $n-$compartment model. For example the parameter vector $\boldsymbol{\alpha}_2$ for the two-compartment model is given as

$$\boldsymbol{\alpha}_2 = [k_a^{(2)} \;\; k_b^{(2)} \;\; k_{out}^{(2)} \;\; v_e^{(2)} \;\; v_p^{(2)}]. \tag{15}$$

The ICG measurements in equation (14) are collected at discrete time instances, $t = kT$, $k = 0, 1, ...,$ where T is the sampling period. Therefore, the continuous model described in equations (14) has to be discretized. To simplify our notation, we shall use $\mathbf{C}(k) = \mathbf{C}(kT)$ and $\mathbf{m}(k) = \mathbf{m}(kT)$. The discrete KF system and the measurement models are given as follows:

$$\mathbf{C}(k+1) = \kappa_d(\boldsymbol{\alpha}_n)\mathbf{C}(k) + \boldsymbol{\omega}(k)$$

$$\mathbf{m}(k) = V_d(\alpha_n)\mathbf{C}(k) + \eta(k), \tag{16}$$

where $\kappa_d(\alpha_n) = e^{\kappa(\alpha_n)}$ is the discrete-time KF system matrix and $V_d(\alpha_n) = V(\alpha_n)$ is the discrete-time KF measurement matrix. $\omega(k)$ and $\eta(k)$ are zero mean Gaussian white noise processes with covariances matrix Q_d and variance σ_d^2, respectively. Discretization of state-space models can be found in various system theory books, see for example [25].

An explicit form of the discrete KF model for the two-compartment case is given as follows:

$$\begin{bmatrix} C_e(k+1) \\ C_p(k+1) \end{bmatrix} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix} \begin{bmatrix} C_e(k) \\ C_p(k) \end{bmatrix} + \omega(k) \tag{17}$$

$$\mathbf{m}(k) = \begin{bmatrix} v_e^{(2)} & v_p^{(2)} \end{bmatrix} \begin{bmatrix} C_e(k) \\ C_p(k) \end{bmatrix} + \eta(k),$$

where τ_{ij} is the ith row and j^{th} column entry of the system matrix $\kappa_d(\alpha_2)$. Note that the matrix entry τ_{ij} is an exponential function of the parameters $k_a^{(2)}$, $k_b^{(2)}$ and $k_{out}^{(2)}$.

To simplify the estimation process, we shall first estimate the matrix entries, τ_{ij}, of the discrete-time system matrix $\kappa_d(\alpha_n)$ and then compute the pharmacokinetic parameters for each compartmental model.

B. Modeling of ICG Pharmacokinetic Parameters and Concentrations in Extended Kalman Filter Framework

Kalman filter provides a recursive method to estimate the states in state-space models, in which the states are driven by noise, and the measurements are collected in the presence of measurement noise [26], [27], [28]. In the case of non-linear state-space models, the extended Kalman filter linearizes the model around the current state estimate, and then applies the KF to the resulting linear model. The EKF framework is also utilized for the joint estimation of the unknown system and/or measurement parameters and states. In a linear state-space model when both states and system parameters are unknown, the linear state-space model can be regarded as a non-linear model in which the linear system parameters and states are combined to form the new states of the non-linear model. This system is then linearized and solved for the unknown states using the KF estimator [29], [30], [31]. In our problem, our objective is to simultaneously estimate the states, i.e., the ICG concentrations in each compartment, and the system and measurement parameters, i.e., the pharmacokinetic parameters and the volume fractions.

We consider a Taylor series approximation to the non-linear system function at the previous state estimates, and that of the measurement function at the corresponding predicted position. This approach provides a simple and efficient method to handle the non-linearity in the new system and measurement models.

Let $\boldsymbol{\theta}_n(k)$ denote the discrete-time parameter vector of the pharmacokinetic rates and volume fractions at time k. For example, in the two-compartment model, $\boldsymbol{\theta}_2(k)$ is given as

$$\boldsymbol{\theta}_2(k) = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{21} & \tau_{22} & v_e^{(2)} & v_p^{(2)} \end{bmatrix}^T . \tag{18}$$

In the EKF framework, $\boldsymbol{\theta}_n(k)$ is treated as a random process with the following model:

$$\boldsymbol{\theta}_n(k+1) = \boldsymbol{\theta}_n(k) + \varsigma(k), \tag{19}$$

where $\varsigma(k)$ is a zero mean white noise process with covariance matrix S_d.

We append the parameter vector $\boldsymbol{\theta}_n(k+1)$ to the ICG concentration vector $\mathbf{C}(k+1)$ to form the new non-linear state-space model given as

$$\begin{bmatrix} \mathbf{C}(k+1) \\ \boldsymbol{\theta}_n(k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{K}(\boldsymbol{\theta}_n)\mathbf{C}(k) \\ \boldsymbol{\theta}_n(k) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\omega}(k) \\ \varsigma(k) \end{bmatrix} \tag{20}$$

$$\mathbf{m}(k) = \begin{bmatrix} V_d(\boldsymbol{\theta}_n) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{C}(k) \\ \boldsymbol{\theta}_n(k) \end{bmatrix} + \eta(k),$$

where $\mathbf{K}(\theta_n) = \kappa_d(\alpha_n)$.

The choice of Q_d, S_d and σ_d^2 is crucial to the performance EKF estimator. It was shown that if these values are selected less than the actual values, it leads to overconfidence in the accuracy of the estimates of the error covariance matrix. Therefore, these matrices are regarded as tuning parameters and not as the estimates of the true covariance matrices [32].

C. EKF Joint Estimation for ICG Concentrations, Pharmacokinetic Parameters, and Volume Fractions

In this section we will summarize the major steps of the EKF estimator for the joint estimation of ICG concentrations and compartmental model parameters.

Let the subscript $k|t$ denote the estimate at time k given all the measurements up to time t. Then the 1-step ahead prediction of the ICG concentrations and the compartmental model parameters are given as follows:

$$\begin{bmatrix} \hat{\mathbf{C}} \\ \hat{\boldsymbol{\theta}}_n \end{bmatrix}_{k|k-1} = \begin{bmatrix} \mathbf{K}(\hat{\boldsymbol{\theta}}_n)\hat{\mathbf{C}} \\ \hat{\boldsymbol{\theta}}_n \end{bmatrix}_{k-1|k-1} . \tag{21}$$

For the two-compartment model, Equation (21) becomes

$$\begin{bmatrix} \hat{C}_e \\ \hat{C}_p \\ \hat{\boldsymbol{\theta}}_2 \end{bmatrix}_{k|k-1} = \begin{bmatrix} \widehat{\tau_{11}}\hat{C}_e + \widehat{\tau_{12}}\hat{C}_p \\ \widehat{\tau_{21}}\hat{C}_e + \widehat{\tau_{22}}\hat{C}_p \\ \hat{\boldsymbol{\theta}}_2 \end{bmatrix}_{k-1|k-1} . \tag{22}$$

The error covariance matrix, $P_{k|k-1}$, of the 1-step ahead predictions is given as follows:

$$P_{k|k-1} = J_{k-1}P_{k-1|k-1}J_{k-1}^T + \begin{bmatrix} Q_d & 0 \\ 0 & S_d \end{bmatrix}, \tag{23}$$

where J_k is the Jacobian of the non-linear EKF system function at time k. Explicitly, it is given as:

$$J_k = \begin{bmatrix} \mathbf{K}(\hat{\boldsymbol{\theta}}_n) & \frac{\partial}{\partial \boldsymbol{\theta}_n}[\mathbf{K}(\hat{\boldsymbol{\theta}}_n)\hat{\mathbf{C}}] \\ \mathbf{0} & \mathbf{I} \end{bmatrix}_{k|k} , \tag{24}$$

where **0** and **I** denote zero and identity matrices, respectively. The Jacobian matrix for the two-compartment model becomes

$$J_k = \begin{bmatrix} \begin{pmatrix} \widehat{\tau_{11}} & \widehat{\tau_{12}} \\ \widehat{\tau_{21}} & \widehat{\tau_{22}} \end{pmatrix} & \begin{pmatrix} \hat{C}_e & \hat{C}_p & 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{C}_e & \hat{C}_p & 0 & 0 \end{pmatrix} \\ \mathbf{0}_{(6\times2)} & \mathbf{I}_{(6\times6)} \end{bmatrix}_{k|k} , \tag{25}$$

where $\mathbf{0}_{(6\times 2)}$ is a 6×2 zero matrix, and $\mathbf{I}_{(6\times 6)}$ is a 6×6 identity matrix.

The 1-step ahead predictions are updated to the kth-step estimates by means of the Kalman gain matrix which is given by

$$G_k = P_{k|k-1}\Lambda^T[\Lambda P_{k|k-1}\Lambda^T + \sigma_k^2]^{-1}, \tag{26}$$

where Λ is the following vector

$$\begin{bmatrix} V_d(\hat{\boldsymbol{\theta}}) & \frac{\partial}{\partial \boldsymbol{\theta}}[V_d(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}] \end{bmatrix}_{k|k-1}. \tag{27}$$

For the two-compartment model Λ vector becomes

$$\begin{bmatrix} \hat{v_e}^{(2)} & \hat{v_p}^{(2)} & 0 & 0 & 0 & 0 & \hat{C_e} & \hat{C_p} \end{bmatrix}_{k|k-1}. \tag{28}$$

The kth-step estimate of the concentrations and the parameters are obtained recursively using

$$\begin{bmatrix} \hat{\mathbf{C}} \\ \hat{\boldsymbol{\theta}} \end{bmatrix}_{k|k} = \begin{bmatrix} \hat{\mathbf{C}} \\ \hat{\boldsymbol{\theta}} \end{bmatrix}_{k|k-1} + G_k(\mathbf{m}(k) - [V_d(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}]_{k|k-1}). \tag{29}$$

For the two-compartment case, the kth-step estimate of the concentrations and the parameters is

$$\begin{bmatrix} \hat{C_e} \\ \hat{C_p} \\ \hat{\boldsymbol{\theta}}_{(2)} \end{bmatrix}_{k|k} = \begin{bmatrix} \hat{C_e} \\ \hat{C_p} \\ \hat{\boldsymbol{\theta}}_2 \end{bmatrix}_{k|k-1} + G_k(\mathbf{m}(k) - (\hat{v_e}^{(2)}\hat{C_e} - \hat{v_p}^{(2)}\hat{C_p})_{k|k-1})). \tag{30}$$

The error covariance matrix, $P_{k|k}$, of the kth-step estimates is updated as

$$P_{k|k} = [I - G_k\Lambda]P_{k|k-1}, \tag{31}$$

where $\mathbf{I}$ is the identity matrix.

The initialization of the ICG concentrations, pharmacokinetic parameters, and the volume fractions plays an important role in the performance of the EKF algorithm. Theoretically, the state estimates can be initialized at the expected value of the ICG concentrations, i.e. $E[\mathbf{C}(0)]$.

One approach to the initialization of the parameters is to utilize the state-space presentation given in equation (16). Since $E(\mathbf{m}(0)) = V_d(\boldsymbol{\theta}_n(0))E[\mathbf{C}(0)]$, $\mathbf{m}(0) - V_d(\boldsymbol{\theta}_n(0))E[\mathbf{C}(0)]$ is a zero mean random variable. If we express the variance of the measurement $\mathbf{m}(0)$ in terms of the variance of $\mathbf{C}(0)$ using the measurement model in equation (16), and solve for $\boldsymbol{\theta}_n$, we get the estimate $\hat{\boldsymbol{\theta}}_n(0)$ as the most appropriate value for initialization. The details of the selection of the initial values for the parameters can be found in [27].

The initialization of the error covariance matrix is also important for the performance of the EKF. The error covariance matrix is the matrix which provides information about the error bounds for the estimates. Theoretically, the initial error covariance matrix is a diagonal matrix where the diagonal entries are the initial estimates of the variance of concentrations and pharmacokinetic parameters, i.e.

$$P_{0|0} = \begin{bmatrix} Cov(\mathbf{C}(0)) & 0 \\ 0 & S_d \end{bmatrix}. \tag{32}$$

D. Compartmental Model Order Selection

We adopted Bayesian information criteria (BIC) for the optimal model order selection. BIC is a well known information theoretic criteria, in which the optimal model order is selected by minimizing a cost function to avoid overfitting. The cost function depends on the number of observations, number of unknown parameters to be estimated and the maximum likelihood function. A detailed discussion on BIC can be found in [33], [34], [35].

In order to calculate the BIC for different compartmental models, we first derived a likelihood function for the extended Kalman filter. The derivation is based on the maximum likelihood estimation of the parameters in the Kalman filtering framework given as in [36], [37]. We then modified this likelihood function for the extended Kalman filter estimator for the joint estimation of compartmental model parameters and concentrations. The details of the derivation is provided in the appendix.

IV. EXPERIMENTAL RESULTS - ICG PHARMACOKINETICS IN FISCHER RAT DATA

We applied the proposed EKF framework to the pharmacokinetic analysis of ICG data obtained from four Fischer rats with adenocarcinoma. R3230ac adenocarcinoma cells were injected below the skin into four Fischer rats 3 weeks prior to measurements. The tumor size for the rats ranges in diameter from 5 to 30 mm. The ICG concentration data was collected with an MRI-NIR imager. The configuration of the apparatus, the data collection procedure, and the details of the experimental approach have been reported in [9], [38], [39].

Figure 4 presents the ICG concentrations (μM) from four different rats. Tumors in Rat 1 and 2 are classified as necrotic because of their low tissue oxy-hemoglobin, low total hemoglobin, and low gadolinium-diethylene-triamine penta-acetic acid (Gd-DTPA) enhancement levels. Tumors in Rat 3 and 4 are classified as edematous due to their high water content [40]. It can be observed from Figure 4 that the necrotic cases display low peak ICG concentration values and slowly rising slopes unlike the edematous cases with high peak values and sharp rising slopes.

We estimated the pharmacokinetic rates for the four-, three- and two-compartment models. The results are given in Tables I, II, and III, respectively. The error bounds on the estimates are derived from the covariance matrix of the EKF estimator. The estimated pharmacokinetic rates for all compartmental models indicate that the exchange rates between the capillary and the adjacent compartment (ISS or EES), k_a^n, k_b^n, $n = 2, 3, 4$, are significantly different for the necrotic and edematous tissue. We observe that for the four- and three-compartment models, the estimated exchange rates between the ISS and parenchymal cell compartments, k_c^n, k_d^n, $n = 3, 4$, are comparable. Similarly, the estimated rate of drainage out of the plasma, k_{out}^n, $n = 2, 3, 4$, are consistent for all models.

Based on the model parameter estimates, we computed the BIC values for each rat data to reveal overfitting. The BIC values and the number of unknown parameters for each rat data are tabulated in Table IV. The BIC suggests that the two-compartment model is sufficient for all four measurement sets.

We further analyze the goodness-of-fit of the compartmental models by means of the residual analysis. The basic idea of the residual analysis is to compare the actual measurements $\mathbf{m}(k)$ with its 1-step ahead predictions, $\hat{\mathbf{m}}(k)_{k|k-1}$, based on the estimated parameters. A detailed discussion on residual analysis can be found in [26], [41]. The mean and variance of the residual error for the four-, three- and two-compartmental models are tabulated in Table V. To normalize the error with respect to the magnitude of the actual measurements, we calculated the signal-to-noise ratio (SNR) using the median value of the measurements and the mean of the residual errors for each rat data. As seen from the results in Table VI, the SNR values are higher for the two-compartment case for all data sets. These results show that the two-compartment model provides the minimum bias and the best statistical efficiency. Figure 5 shows the measured total concentration data and its 1-step ahead prediction based on the two-compartment model for each rat data. Clearly, there is a good agreement between the actual and the estimated measurements.

Based on the BIC and residual analysis, we conclude that the two-compartment model provides the best statistical fit for the rat data and investigate the estimated model parameters in more detail.

In the two-compartment model, the rate of leakage into the EES from the capillary, $k_a^{(2)}$, range from 0.0106 to 0.0777 sec^{-1} and the rate of drainage out of the EES and into the capillary, $k_b^{(2)}$, range from 0.0247 to 0.0840 sec^{-1}. Note that the permeability rates for the necrotic cases are lower than the ones observed for the edematous cases. Additionally, the estimated values for the pharmacokinetic rates are much higher than the normal tissue values due to the increased leakiness of the blood vessels around the tumor region [9], [42]. The estimated plasma volume fractions agrees with the values reported earlier [9], and the values presented in the literature [43], [44]. These results confirm that $v_p^{(2)}$ can be significantly large in tumors and that its magnitude varies with respect to the stage of the tumor [24]. The estimated values of EES volume fraction, $v_e^{(2)}$, range from 0.218 to 0.53, in agreement with 0.2 to 0.5 range reported earlier [23]. Note that these results are valid only for the ICG pharmacokinetics in tumor cells R3230ac, adenocarcinoma and may not be generalized for other types of contrast agents or tumor types.

Figure 6 shows the estimated ICG concentrations in the plasma and the EES compartments for the two-compartment model for Rats 1 to 4. Note that initial estimates of concentrations are noisy due to the limited data used in the recursive EKF estimation. This can be improved by Kalman backward smoothing [45]. The peak values of the plasma concentration, C_p, range from 2.72 μM to 4.28 μM. The absolute value of the concentrations may not be very useful. However, concentration of ICG in a compartment relative to the one in another compartment may provide useful information. We consider the ratio of the peak concentrations in the plasma and EES as a potential parameter to discriminate different tumors. The peak C_p/C_e ratio for Rats 1 to 4 is 0.551, 0.593, 0.787, 1.151, respectively. This ratio is higher in edematous cases consistent with the fact that ICG-albumin leaks more into the EES in edematous tumors. Additionally, the ICG concentration in plasma decays faster than the ICG concentration in EES due to its elimination through the liver and kidneys.

V. CONCLUSION

In this paper we present three different compartmental models, an extended Kalman filtering framework for the modeling, and estimation of ICG pharmacokinetics in cancerous tumors based on NIR measurements. Additionally, we introduce information theoretic criteria and residual analysis for model selection and statistical validation. Proposed compartmental models are fit to data obtained from Fischer rats with adenocarcinoma cells. The pharmacokinetic rates and volume fractions are estimated for all models. The estimated rates for all compartmental models indicate that the exchange rates between the capillary and the adjacent compartment (ISS or EES) are significantly larger for the edematous tissue as compared to the necrotic cases. Based on the BIC and residual analysis, we concluded that the two-compartment model provides the best statistical fit for the rat data and ICG pharmacokinetics. Parameters of this model indicate that the permeability rates are higher for edematous cases as compared to the necrotic tumors. Additionally, we estimated the ICG concentrations in different compartments. The concentrations in different compartments may provide additional parameters for tissue characterization.

While our study indicates that two-compartment provides the best fit for the ICG pharmacokinetics, the four or three-compartment models may be advantageous for modeling the pharmacokinetics of functionalized optical contrast agents that actively accumulate or activate in diseased tissue [46], [47], [48]. In the near future, we plan to investigate three and four-compartmental models in the EKF framework for these optical agents collected from animals, and the ICG data collected from human subjects.

VI. ACKNOWLEDGEMENT

The authors would like to thank Dr. Bruce Tromberg and Dr. David Cuccia for providing the rat data used in our study.

APPENDIX

The cost function for BIC is given as

$$\phi_{BIC}(p) = p \ln N - 2 \ln L(\boldsymbol{\theta}_p, \mathbf{m}(1), \mathbf{m}(2), \ldots\ldots, \mathbf{m}(N)), \tag{33}$$

where p is the dimension of $\boldsymbol{\theta}_p$, which is related to the number of compartments in the model, N is the data length, and $L(\boldsymbol{\theta}, \mathbf{m}(1), \mathbf{m}(2), \ldots\ldots, \mathbf{m}(N))$ is the likelihood function.

The likelihood function for the EKF is given as

$$L(\boldsymbol{\theta}, \mathbf{m}(1), \mathbf{m}(2), \ldots\ldots, \mathbf{m}(N)) = -\frac{1}{2} \sum_{k=1}^{N} \ln[\det(H_k)] - \frac{1}{2} \sum_{k=1}^{N} A_k^T H_k^{-1} A_k, \tag{34}$$

where the matrix H is defined as:

$$H_k = \Lambda P_{k|k-1} \Lambda^T + \sigma_k^2, \tag{35}$$

and σ_k^2, Λ, and $P_{k|k-1}$ are as defined in Section III.C.

The matrix A is defined as:

$$A_k = \mathbf{m}(k) - [V_d(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}]_{k|k-1}, \tag{36}$$

where $\mathbf{m}(k)$ is the ICG concentration data collected from Fisher rats at time k, and $[V_d(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}]_{k|k-1}$ is the 1-step ahead estimate of the volume fractions and concentrations.

The explicit form of the likelihood function for BIC calculation is given as

$$\begin{aligned} L(\boldsymbol{\theta}, \mathbf{m}(1), \mathbf{m}(2), \ldots\ldots, \mathbf{m}(N)) \quad = \quad & -\frac{1}{2} \sum_{k=1}^{N} \ln[\det(\Lambda P_{k|k-1} \Lambda^T + \sigma_k^2)] \\ & -\frac{1}{2} \sum_{k=1}^{N} [\mathbf{m}(k) - [V_d(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}]_{k|k-1}]^T [\Lambda P_{k|k-1} \Lambda^T + \sigma_k^2]^{-1} [\mathbf{m}(k) - [V_d(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}]_{k|k-1}]. \end{aligned}$$

where all the parameters and matrices are as defined in Section III.C.

REFERENCES

[1] A. Yodh and B. Chance, "Spectroscopy and imaging with diffusing light," Phys. Today Vol. 48-3, pp. 34-40, 1995.

[2] X. Intes and B. Chance, "Non-PET functional imaging techniques: optical," *Radiol. Clin. N Am.*, vol.43, pp.221-234, 2005

[3] D. Boas, D. Brooks, E. Miler, C. DiMarzio, M. Kilmer, R. Gaudette, "Imaging the body with diffuse optical tomography," *IEEE Signal Processing Magazine*, vol. 18, pp. 57-74, 2001.

[4] D. Hawrys and E. Sevick-Muraca, "Developments toward diagnostic breast cancer imaging using Near-Infrared optical measurements and fluorescent contrast agents," *Neoplasia* vol.2, pp. 388-417, 2000.

[5] Furukawa, K., D. H. Crean, T. S. Mang, H. Kato and T. J. Dougherty (1995) "Fluorescence detection of premalignant, malignant, and micrometastatic disease using hexylpyropheophorbide," *Proc. SPIE* vol. 2371, pp. 510-514.

[6] Y. Chen , Q. Liu , P. Huang, S. Hyman, X. Intes, W. Lee, and B. Chance, "Assessment of tumor angiogenesis using fluorescence contrast agents," *Proc. of the SPIE*, vol. 5254, no.1, pp. 296-301, 2003.

[7] A. Becker, G. Schneider, B. Riefke, K. Licha, and W. Semmler, "Localization of near-infrared contrast agents in tumors by intravital microscopy," *Proc. SPIE* vol. 3568, pp. 112-118.

[8] M. Gurfinkel, A. B. Thompson, W. Ralston, T. L. Troy, A. L. Moore, T. A. Moore, J. D. Gust, D. Tatman, J. S. Reynolds, B. Muggenburg, K. Nikula, R. Pandey, R. H. Mayer, D. J. Hawrysz, and E. M. Sevick-Muraca, "Pharmacokinetics of ICG and HPPH-car for the detection of normal and tumor tissue using fluorescence, near-infrared reflectance imaging: a case study," *Photochem. Photobiol.*, vol. 72, pp. 94-102, 2000.

[9] D.J. Cuccia, F. Bevilacqua, A. J. Durkin, S. Merritt, B. J. Tromberg, G. Gulsen, H. Yu, J. Wang, and O. Nalcioglu, "In vivo quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration," *Applied Optics*, vol. 42, no. 1, pp. 2940-2950, June 2003.

[10] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, B. Chance, "*In vivo* continuous-wave optical breast imaging enhanced with Indocyanine Green," *Med. Phys.*, vol. 30-6, pp.1039-1047, June 2003.

[11] D. Hansen, A. Spence, T. Carski, and M. Berger, "Indocyanine green (ICG) staining and demarcation of tumor margins in a rat glioma model," *Surg. Neurol.*, vol. 40, pp. 451-456, 1993.

[12] H. Shinohara, A. Tanaka, T. Kitai, N. Yanabu, T. Inomoto, S. Satoh, Hatano, Y. Yamaoka, and K. Hirao, "Direct measurement of hepatic Indocyanine Green clearance with near-infrared spectroscopy: separate evaluation of uptake and removal," *Hepatology*, vol. 23, pp. 137-144, 1996.

[13] A. ElDeosky, A. Seifalian, M. Cope, D. Delpy, and B. Davidson, "Experimental study of liver dysfunction evaluated by direct Indocyanine green clearance using near infrared spectroscopy," *Br. J. Surg.*, vol. 86, pp. 1005-1011, 1999.

[14] X. Li, B. Beauvoit, R. White, S. Nioka, B. Chance, and A. Yodh, "Tumor localization using fluorescence of Indocyanine Green (ICG) in rat models," *Proc. SPIE*, vol. 2389, pp. 789-797, 1995.

[15] M. S. Yates, C. J. Bowmer, and J. Emmerson, "The plasma clearance of indocyanine green in rats with acute renal failure: effect of dose and route of administration," *Biochem. Pharmacol.*, vol. 32, pp. 3109-3114, 1983.

[16] C. W. Tornoe, "Grey-Box PK/PD Modeling of Insulin," M.S. Thesis, June 28, 2002.

[17] D. H. Anderson, *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1983.

[18] J. A. Jacquez, *Compartmental analysis in biology and medicine, Kinetics of distribution of tracer-labeled materials*, Elsevier Pub. Co., New York, 1972.

[19] C. Cobelli, D. Foster, and G. Toffolo, *Tracer kinetics in biomedical research: from data to model*, Kluwer Academic/Plenum, New York, c2000.

[20] P. Huang, X. Intes, B. Chance, S. Nioka, "Simulation of delivery of indocynanine green injected intravenously into the human subject for breast cancer detection," *Proceedings of SPIE*, vol. 4949, pp. 450-459, 2003.

[21] S. Mordon, J. M. Devoisselle, S. Soulie-Begu, T. Desmettre, "Indocyanine Green: Physicochemical Factors Effecting Its Flourescence in vivo," *Microvascular Research*, vol. 55, pp. 146-152, 1998.

[22] S. Fickweiler, R. M. Szeimies, W. Baumler, P. Steinbach, S. Karrer, A. E. Goetz, C. Abels, F. Hofstadler, M. Landthaler, "Indocyanine Green: Intracellular Uptake and Phototherapeutic effects in vitro," *Jour. Photochem. Photobio. B: Biology*, vol. 38, pp. 178-183, 1997.

[23] P. S. Tofts, DPhil, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp, H. B.W. Larsson, T. Lee, N. A. Mayr, G. J.M. Parker, R. E. Port, J. Taylor, and R. M. Weisskoff, "Estimating Kinetic Parameters From Dynamic Contrast-Enhanced T1-Weighted MRI of a Diffusable Tracer: Standardized Quantities and Symbols," *Jour. Mag. Res. Ima.*, vol. 10, pp. 223-232, 1999.

[24] P. S. Tofts, "Modeling tracer kinetics in dynamic Gd-DTPA MR imaging," *J. Magn. Reson. Imag.*, vol. 7, pp. 91-101, 1997.

[25] C. Chen, *Linear System Theory and Design*, Oxford University Press, New York, 1999.

[26] P. Zarchan, *Fundamentals of Kalman filtering : a practical approach*, American Institute of Aeronautics and Astronautics, Reston, 2000.

[27] C.K.Chui, G. Chen, *Kalman Filtering with real time applications*, Springer, Berlin, 1999.

[28] C. E. Catlin, *Estimation, control, and the discrete Kalman filter*, Springer, New York, 1989.

[29] L. Ljung, "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems," *IEEE Tran. Automa. Control*, vol. AC-24, no. 1, pp. 36-50, Feb 1979.

[30] R. Togneri, and L. Deng, "Joint State and Parameter Estimation for a Target-Directed Nonlinear Dynamic System Model," *IEEE Tran. on Sig. Proc.*, vol. 51, no. 12, pp. 3061-3070, Dec 2003.

[31] L. Nelson, and E. Stear, "The Simultaneous On-Line Estimation of Parameters and States in Linear Systems," *IEEE Tran. on Auto. Control*, vol.21, pp. 94-98, Feb 1976.

[32] B. F. La Scala, and R. R. Bitmead, "Design of an extended Kalman filter frequency tracer," *IEEE Tran. on Sig. Proc*, vol. 44, no. 3, pp. 739-742, March 1996.

[33] G. Schwarz, "Estimating the dimensions of a model," *Annals. of Statistics*, vol. 6, pp. 461-464,1978.

[34] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, Third edition, Prentice Hall, 1994.

[35] H. Akaike, "Likelihood and the Bayes Procedure," Bayesian Statistics, Univ. Press, Valencia, Spain, 1980.

[36] A. Harvey, *Time Series Models*, MIT Press, 1993.

[37] W.M. Sallas and D.A. Harville, "Noninformative Priors and Restricted Maximum Likelihood Estimation in the Kalman Filter", in J.C. Spall (Ed.), Bayesian Analysis of Time Series and Dynamic Models, New York: Marcel-Dekker, Inc., 1988.

[38] F. Bevilacqua, A. J. Berger, A. E. Cerussi, D. Jakubowski, and B. J. Tromberg, "Broadband absorption spectroscopy in turbid media by combined frequency-domain and steady-state methods," *Appl. Opt.*, vol. 39, pp. 6498-6507, 2000.

[39] D. J. Jakubowski, "Development of broadband quantitative tissue optical spectroscopy for the non-invasive characterization of breast disease," Beckman Laser Institute, University of California, Irvine, Irvine, Calif., 2002.

[40] S. Merritt, F. Bevilacqua, A. J. Durkin, D. J. Cuccia, R. Lanning, B. J. Tromberg, G. Gulsen, H. Yu, J. Wang, and O. Nalcioglu, "Monitoring tumor physiology using near-infrared spectroscopy and MRI coregistration," *Appl. Opt.*, vol. 42, pp. 2951-2959, 2003.

[41] L.M. Gray, L.D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, Boston, 2004.

[42] M. Y. Su, A. Muhler, X. Lao, and O. Nalcioglu, "Tumor characterization with dynamic contrast-enhanced MRI using MR contrast agents of various molecular weights," *Magn. Reson. Med.*, vol. 39, pp. 259-269, 1998.

[43] J. B. Fishkin, O. Coquoz, E. Anderson, M. Brenner, and B. J. Tromberg, "Frequency-domain photon migration measurements of normal and malignant tissue optical properties in a human subject," *Appl. Opt.*, vol. 36, pp. 10-20, 1997.

[44] D. L. Buckley, "Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T1-weighted MRI," *Magn. Reson. Med.*, vol. 47, pp. 601-606, 2002.

[45] A. Gelb, *Applied Optimal Estimation*, M.I.T. Press, Cambridge, 1989.

[46] K. Licha, "Contrast agents for optical imaging," *Topics in Current Chemistry*, vol.222, pp.1-29, 2002.

[47] Y. Chen, G. Zheng, Z. Zhang, D. Blessington, M. Zhang, H. Li, "Metabolism Enhanced Tumor Localization by Fluorescence Imaging: In Vivo Animal Studies," *Optics Letters*, vol. 28, pp. 2070-2072, 2003.

[48] R. Weissleder, C. H. Tung, U. Mahmood, A. Bogdanov, "In vivo imaging with protease-activated near-infrared fluorescent probes," *Nat. Biotech.*, vol. 17, pp. 375-378, 1999.

TABLE I

FOUR-COMPARTMENT MODEL:ESTIMATED PHARMACOKINETIC PARAMETERS USING EKF ALGORITHM

	$k_a^{(4)}$ ($sec^{-1}10^{-2}$)	$k_b^{(4)}$ ($sec^{-1}10^{-2}$)	$k_c^{(4)}$ ($sec^{-1}10^{-2}$)	$k_d^{(4)}$ ($sec^{-1}10^{-2}$)	$k_e^{(4)}$ ($sec^{-1}10^{-2}$)	$k_f^{(4)}$ ($sec^{-1}10^{-2}$)	$k_{out}^{(4)}$ ($sec^{-1}10^{-3}$)
Rat 1 (Necrotic)	1.45±0.013	1.22±0.019	1.86±0.017	2.02±0.026	2.74±0.041	2.41±0.051	4.05±0.059
Rat 2 (Necrotic)	3.48±0.048	2.77±0.034	4.28±0.048	4.33±0.040	2.98±0.048	3.03±0.061	4.76±0.062
Rat 3 (Edematous)	4.94±0.052	5.16±0.067	4.22±0.052	4.13±0.067	4.14±0.070	4.27±0.078	5.39±0.085
Rat 4 (Edematous)	5.25±0.053	5.31±0.063	5.07±0.068	5.22±0.063	4.43±0.075	4.03±0.072	3.85±0.056

TABLE II

THREE-COMPARTMENT MODEL:ESTIMATED PHARMACOKINETIC PARAMETERS USING EKF ALGORITHM

	$k_a^{(3)}$ ($sec^{-1}10^{-2}$)	$k_b^{(3)}$ ($sec^{-1}10^{-2}$)	$k_c^{(3)}$ ($sec^{-1}10^{-2}$)	$k_d^{(3)}$ ($sec^{-1}10^{-2}$)	$k_{out}^{(3)}$ ($sec^{-1}10^{-3}$)
Rat 1 (Necrotic)	1.93±0.061	1.28±0.049	1.82±0.032	2.02±0.041	3.89±0.052
Rat 2 (Necrotic)	4.41±0.074	2.48±0.067	4.87±0.066	5.03±0.057	5.45±0.071
Rat 3 (Edematous)	4.71±0.085	3.88±0.077	4.95±0.059	4.68±0.050	4.42±0.040
Rat 4 (Edematous)	5.29±0.091	6.48±0.096	4.48±0.062	4.20±0.048	5.01±0.055

TABLE III

TWO-COMPARTMENT MODEL:ESTIMATED PHARMACOKINETIC PARAMETERS AND VOLUME FRACTIONS USING EKF ALGORITHM

	$k_a^{(2)}$ ($sec^{-1}10^{-2}$)	$k_b^{(2)}$ ($sec^{-1}10^{-2}$)	$k_{out}^{(2)}$ ($sec^{-1}10^{-3}$)	$v_e^{(2)}$ (10^{-2})	$v_p^{(2)}$ (10^{-2})
Rat 1 (Necrotic)	2.47±0.043	1.06±0.052	4.61±0.073	21.8±1.92	1.41±0.053
Rat 2 (Necrotic)	3.54±0.082	2.98±0.086	4.83±0.092	25.4±3.49	2.42±0.088
Rat 3 (Edematous)	6.90±0.101	4.93±0.072	3.95±0.048	30.4±2.81	4.84±0.120
Rat 4 (Edematous)	8.40±0.114	7.77±0.091	4.02±0.068	53.0±4.73	7.03±0.321

TABLE IV

TEST FOR THE MODEL ORDER SELECTION FOR THREE DIFFERENT COMPARTMENTAL MODELS FOR FOUR DIFFERENT DATA SETS

		Rat1	**Rat2**	**Rat3**	**Rat4**
Model	p	$\phi_{BIC}(p)$	$\phi_{BIC}(p)$	$\phi_{BIC}(p)$	$\phi_{BIC}(p)$
Two-compartment Model	7	-178.242	-198.367	-202.81	-172.098
Three-compartment Model	11	-71.615	-83.849	-92.182	-63.912
Four-compartment Model	15	-39.719	-45.121	-56.340	-30.023

TABLE V

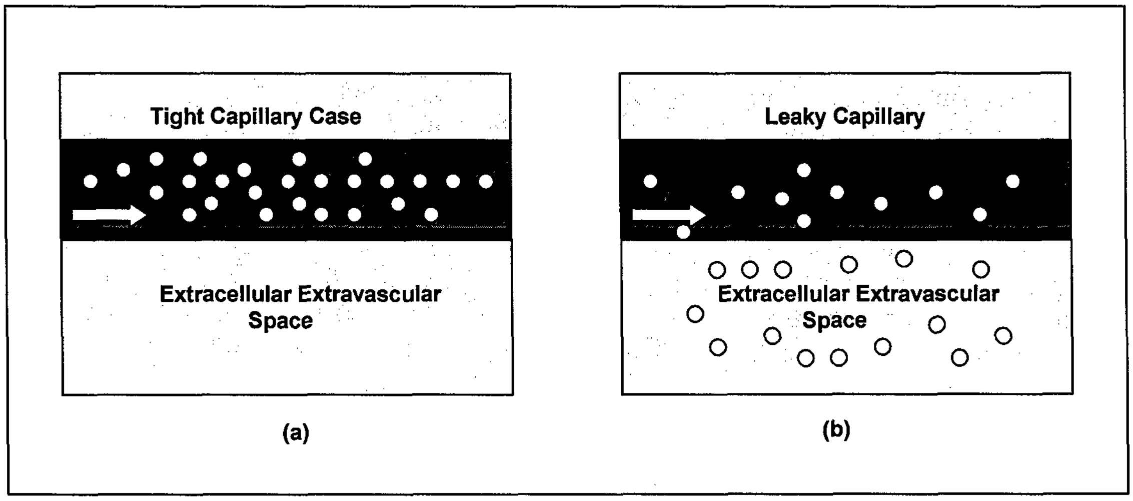
THE MEAN AND VARIANCE OF THE ERROR BETWEEN THE ESTIMATES AND MEASUREMENTS

	Four-compartment		**Three-compartment**		**Two-compartment**	
	Mean	Variance	Mean	Variance	Mean	Variance
Rat1	0.0987	7.641e-004	0.0605	4.732e-004	0.0072	2.567e-005
Rat2	0.1043	9.152e-004	0.0767	3.017e-004	0.0057	4.829e-005
Rat3	0.1204	8.905e-004	0.0883	4.921e-004	0.0041	3.021e-005
Rat4	0.0904	5.977e-004	0.0589	6.839e-004	0.0076	8.618e-005

TABLE VI

SNR VALUES FOR THREE DIFFERENT COMPARTMENTAL MODELS FOR FOUR DIFFERENT DATA SETS

	Rat1	**Rat2**	**Rat3**	**Rat4**
Model	SNR (dB)	SNR (dB)	SNR (dB)	SNR (dB)
Two-compartment Model	73.2	68.1	108.3	107.9
Three-compartment Model	30.7	36.1	23.9	47.0
Four-compartment Model	20.8	29.9	27.7	18.4


Tight Capillary Case
Extracellular Extravascular Space
Leaky Capillary
Extracellular Extravascular Space

(a) (b)

Fig. 1

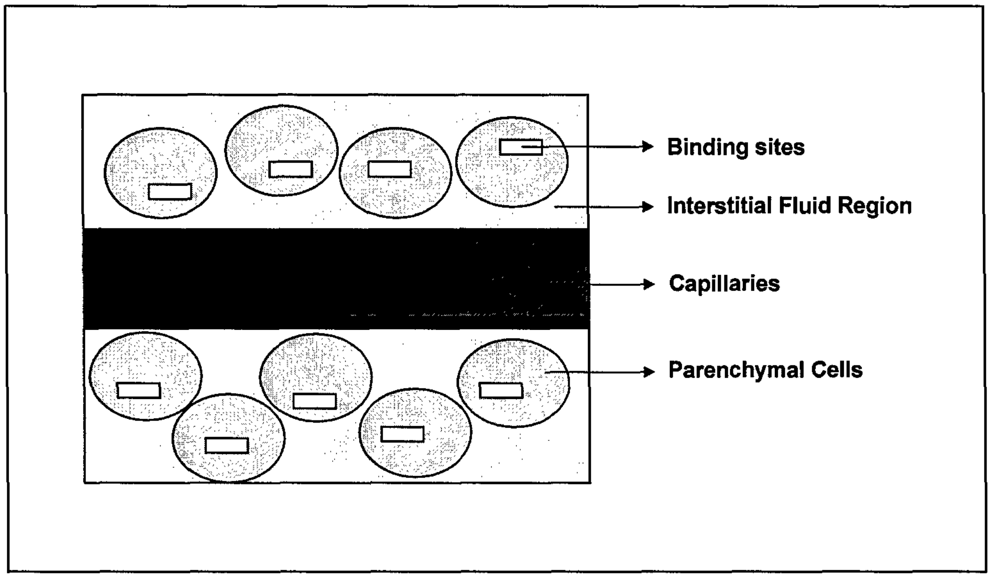

Binding sites
Interstitial Fluid Region
Capillaries
Parenchymal Cells

Fig. 2

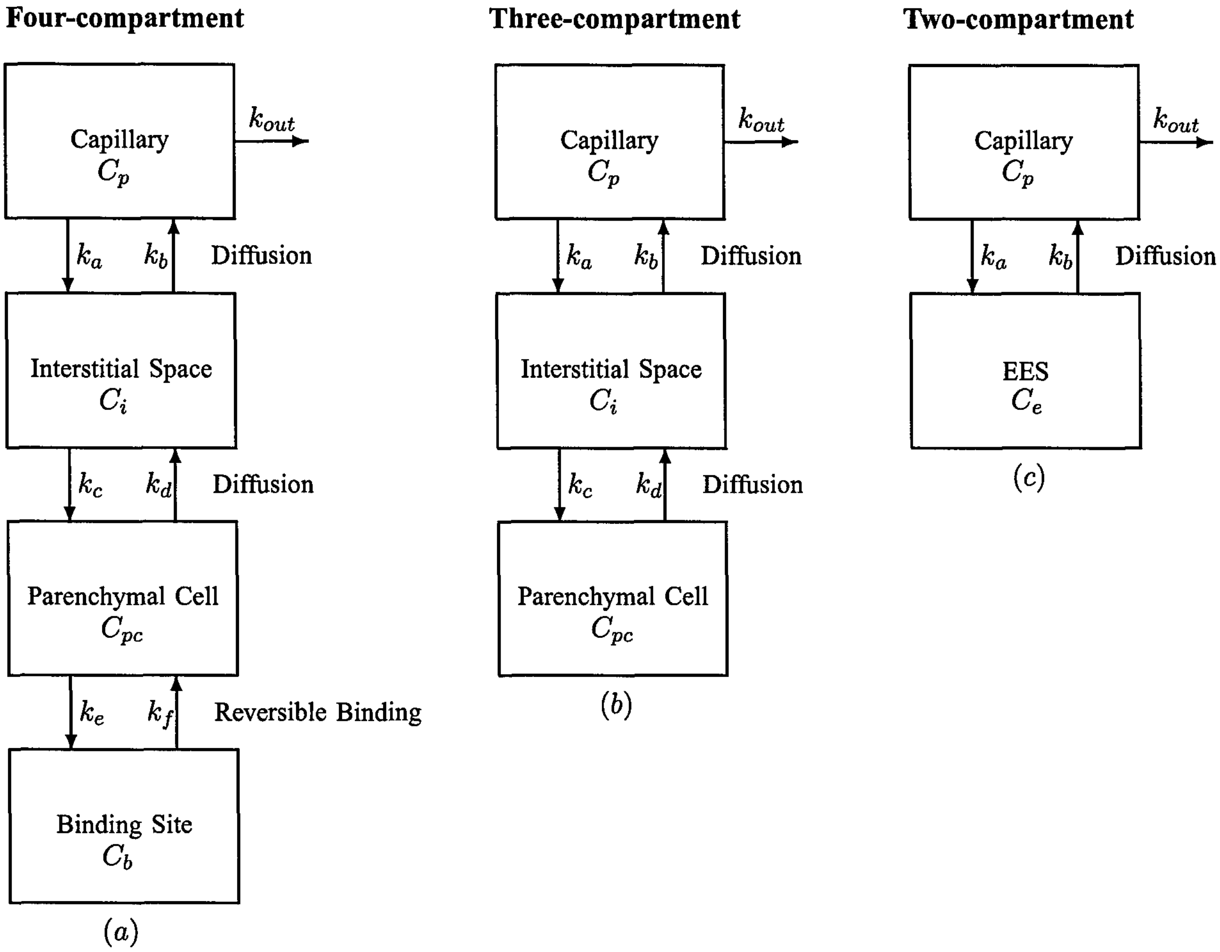

Four-compartment
Three-compartment
Two-compartment
Capillary
C_p
k_{out}
k_a
k_b
Diffusion
Interstitial Space
C_i
k_c
k_d
Diffusion
Parenchymal Cell
C_{pc}
k_e
k_f
Reversible Binding
Binding Site
C_b
(a)
(b)
EES
C_e
(c)

Fig. 3

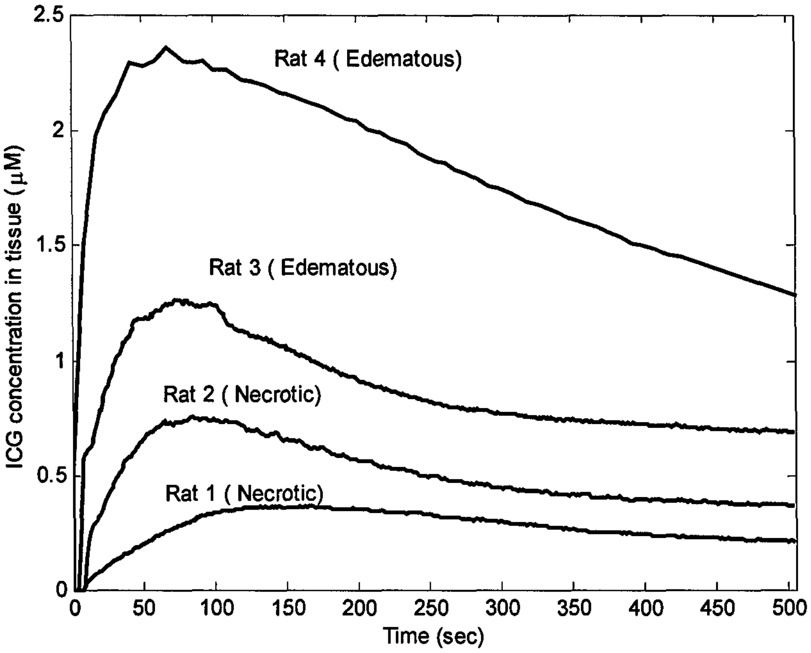
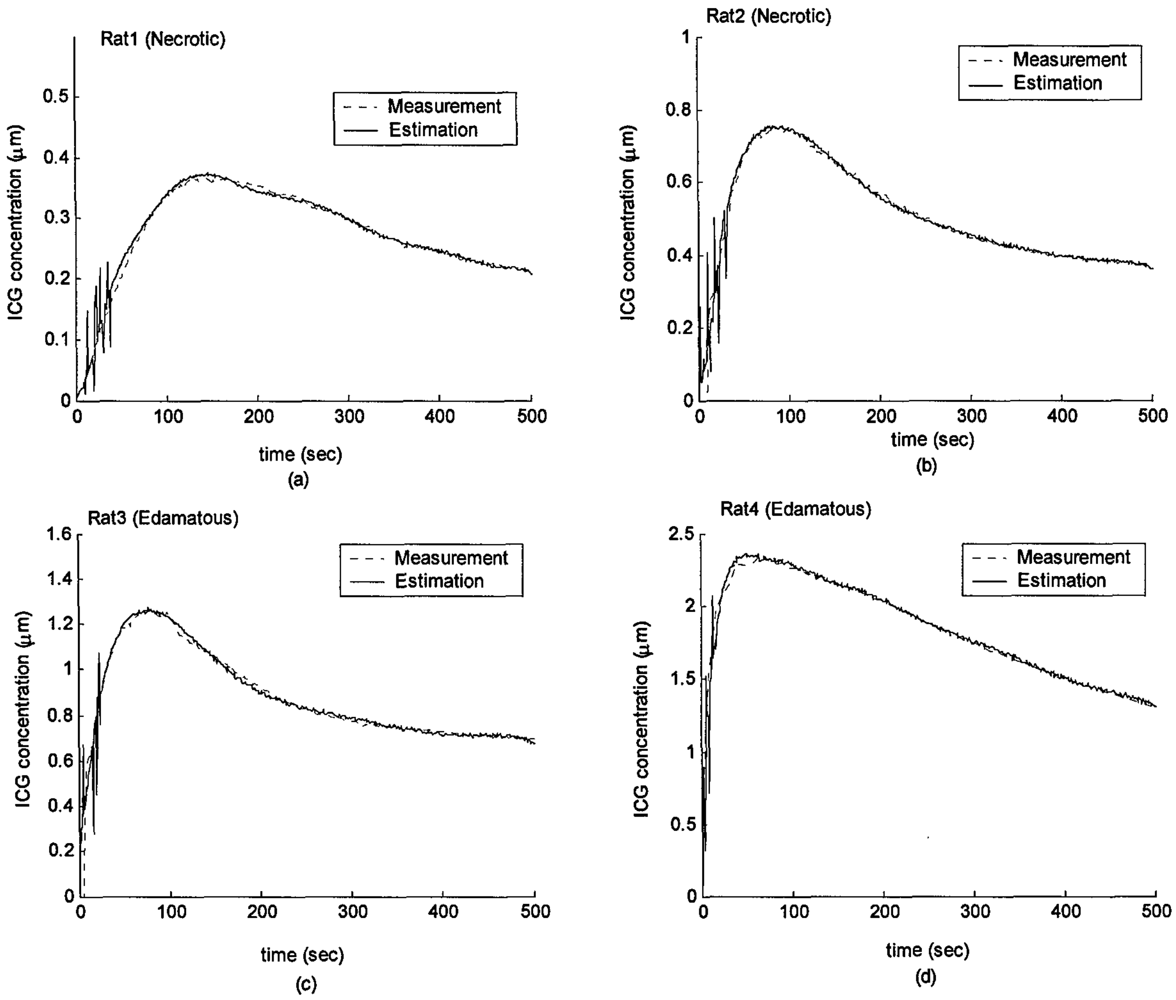

2.5
2
1.5
1
0.5
0
ICG concentration in tissue (μM)
0
50
100
150
200
250
300
350
400
450
500
Time (sec)
Rat 4 (Edematous)
Rat 3 (Edematous)
Rat 2 (Necrotic)
Rat 1 (Necrotic)

Fig. 4


Rat1 (Necrotic)
Rat2 (Necrotic)
Rat3 (Edamatous)
Rat4 (Edamatous)
Measurement
Estimation
ICG concentration (μm)
time (sec)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.8
1
1.2
1.4
1.6
1.5
2
2.5
100
200
300
400
500

(a)

(b)

(c)

(d)

Fig.5

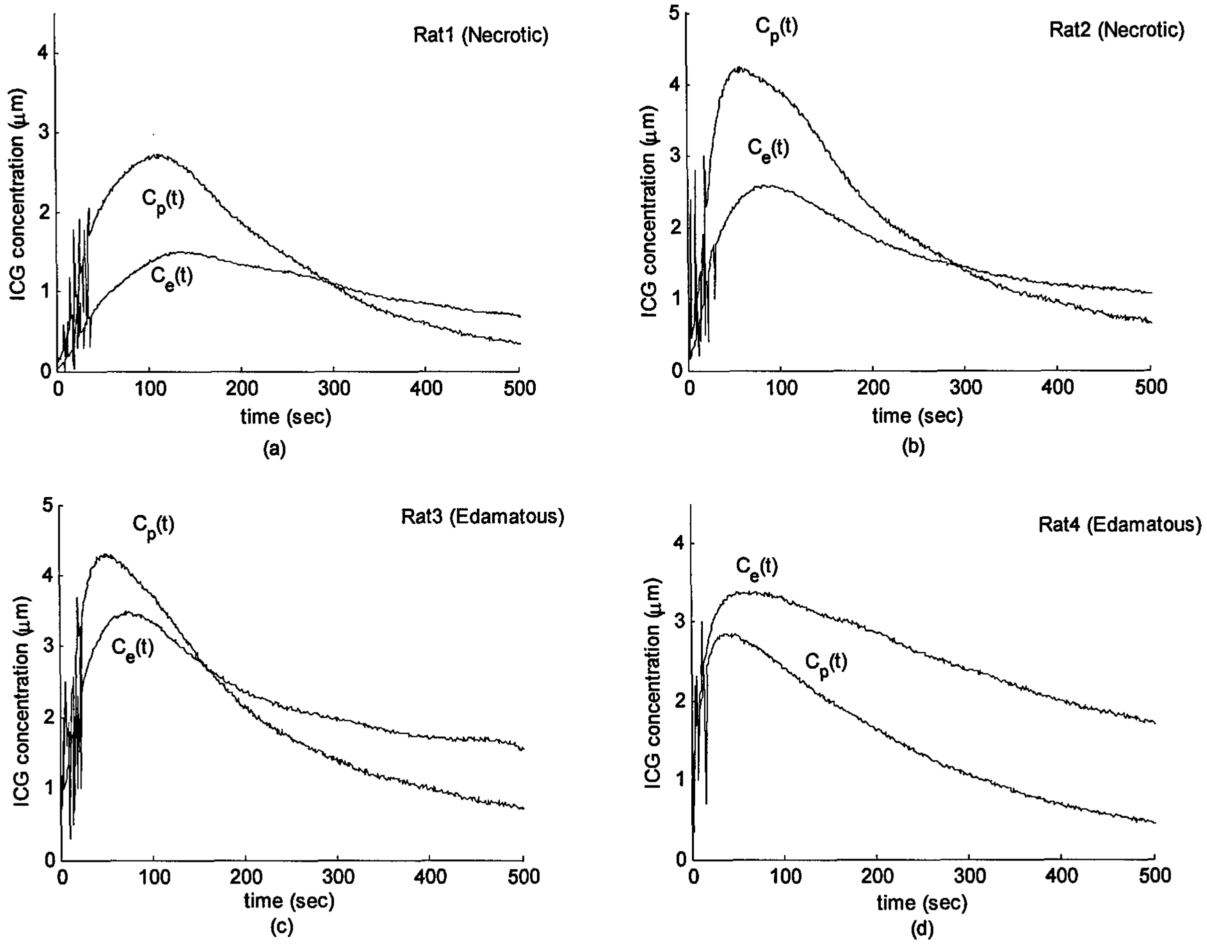

Rat1 (Necrotic)
C_p(t)
C_e(t)
ICG concentration (μm)
0
1
2
3
4
0
100
200
300
400
500
time (sec)
(a)
Rat2 (Necrotic)
C_p(t)
C_e(t)
ICG concentration (μm)
0
1
2
3
4
5
0
100
200
300
400
500
time (sec)
(b)
Rat3 (Edamatous)
C_p(t)
C_e(t)
ICG concentration (μm)
0
1
2
3
4
5
0
100
200
300
400
500
time (sec)
(c)
Rat4 (Edamatous)
C_e(t)
C_p(t)
ICG concentration (μm)
0
1
2
3
4
0
100
200
300
400
500
time (sec)
(d)

Fig. 6

FIGURE CAPTIONS

Fig. 1. A simple illustration of the capillary extracapillary structure.

Fig. 2. An illustration of the ICG flow (a) in tight capillary of normal vessel, (b) in permeable capillary of tumor tissue.

Fig. 3. Block diagram of (a) the four-compartment, (b) the three-compartment, and (c) the two-compartment models for ICG pharmacokinetics.

Fig. 4. ICG concentrations measured in tissue for four different rats.

Fig. 5. ICG concentration measurement data and 1-step prediction of the measurements for four different rats.

Fig. 6. ICG concentrations in plasma, $C_p(t)$ and EES, $C_e(t)$, for four different rats. (a) Rat1, (b) Rat2, (c) Rat3, and (d) Rat4.